

09/832.929

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:21:59 ; Search time 85 Seconds
(without alignments)
1092.412 Million cell updates/sec

Title: US-09-832-929-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHFKDLGEBNFK.....TCFAEGKLVAAQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	3103	100.0	585	10	Mature human serum
2	3103	100.0	585	11	Human albumin (HA)
3	3103	100.0	585	11	Human serum albumin
4	3103	100.0	585	11	Human serum albumin
5	3103	100.0	585	18	Human serum albumin
6	3103	100.0	585	21	HSA protein sequen
7	3103	100.0	585	21	Amino acid sequenc
8	3103	100.0	585	21	Yeast codon-biased
9	3103	100.0	585	22	Human mature album
					Human albumin (HA)

10	3103	100.0	585	22	AA52567	Mature human serum
11	3103	100.0	585	22	AAE13129	Human albumin (HA)
12	3103	100.0	585	22	AAE12403	Human albumin (HA)
13	3103	100.0	585	22	AAE08578	Human serum albumi
14	3103	100.0	585	23	ABG71291	Glycosylated prote
15	3103	100.0	585	23	ABG63321	Human serum albumi
16	3103	100.0	585	23	ABJ00986	B lymphocyte strimu
17	3103	100.0	585	23	ABG33847	Human B lymphocyte
18	3103	100.0	585	23	AAU75220	Mature form of hum
19	3103	100.0	609	21	AAAB36542	Recombinant human
20	3103	100.0	609	21	AAAB36549	Recombinant human
21	3103	100.0	609	21	AAAY78147	pre human serum al
22	3103	100.0	609	24	ABU57252	Human serum albumi
23	3103	100.0	609	24	ABU57253	Human serum albumi
24	3103	100.0	610	14	AAAR39510	Chimeric human ser
25	3103	100.0	616	24	AAE30916	Val8-GJP-1-human s
26	3103	100.0	624	24	AAE30919	Human serum albumi
27	3103	100.0	631	24	AAE30917	Val8-GLP-1-linker-
28	3103	100.0	640	24	AAE30918	Gly8-GLU22-GLP-1-C
29	3103	100.0	640	24	AAE30920	Exendin-4-linker-h
30	3103	100.0	670	21	AAAB36543	Recombinant human
31	3103	100.0	670	21	AAAB36550	Recombinant human
32	3103	100.0	783	14	AAAR39473	Prepro-HSA-G-CSF c
33	3103	100.0	787	14	AAAR39477	G-CSF-(GLY)4-HSA c
34	3103	100.0	853	14	AAAR39472	HSA-vWF(470-713) f
35	3099	99.9	585	10	AAAP93344	Sequence of mature
36	3099	99.9	585	19	AAAW59841	Mature protein of
37	3099	99.9	608	17	AAAR62229	Human serum albumi
38	3099	99.9	609	4	AAAP30089	Sequence of human
39	3099	99.9	609	17	AAAR62232	Human serum albumi
40	3099	99.9	609	17	AAAR94572	Cancer metastasis
41	3099	99.9	609	17	AAAR88913	Human serum albumi
42	3099	99.9	609	19	AAAW48095	Human serum albumi
43	3099	99.9	609	20	AAAY66994	Human albumin. Ho
44	3099	99.9	609	22	AAAB04148	Myosin light chain
45	3099	99.9	609	23	ABG32802	Human serum albumi

ALIGNMENTS

RESULT 1	AAAP90388	standard; protein; 585 AA.
ID	AAAP90388	standard; protein; 585 AA.
XX	AAAP90388;	
AC	AAAP90388;	
XX	25-MAR-2003 (updated)	
DT	01-NOV-1989 (first entry)	
XX	Mature human serum albumin polypeptide.	
DE	Mature human serum albumin polypeptide.	
XX	Human serum albumin; mature protein; new polypeptides;	
KW	plasma expanders.	
XX	Homo sapiens (Human).	
OS	EP322034-A.	
PN	28-JUN-1989.	
XX	25-OCT-1988; 88EP-0310000.	
PD	30-OCT-1987; 87GB-0025529.	
XX	(DELZ) DELTA BIOTECHNOLOGY LTD.	
PA	Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;	
XX	WPI; 1989-186464/26.	
DR	N-PSDB; AAAN90128.	
XX	New N-terminal fragments of human serum albumin	
PT		

```

PT - esp. useful as blood plasma expanders.
XX
PS Disclosure; fig 2; 20pp; English.
XX
CC Mature protein of human serum albumin (see corresp. AAN90128).
CC Used to make new N-terminal fragments which are used as plasma
CC expanders, or as substitutes for HSA or BSA, in tissue culture
CC media.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPRLVRPEV 120
QY 121 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
DB 121 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
QY 181 KLDELDRDEGKASSAKORLUKASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKORLUKASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTTECHGDLLECADRADLAKYICENODSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTTECHGDLLECADRADLAKYICENODSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYBYARRHPDYSVVLRLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYBYARRHPDYSVVLRLRLAKTYETTTLEKC 360
QY 361 CAADAPHECYAKVDFEPKPLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADAPHECYAKVDFEPKPLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVVLNQLCVLHKTPTVSDRVTCKCTES 480
DB 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVVLNQLCVLHKTPTVSDRVTCKCTES 480
QY 481 LVNRRPCFSALVEDETYVPKFNAPETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
DB 481 LVNRRPCFSALVEDETYVPKFNAPETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

RESULT 2
AAR05318
ID AAR05318 standard; protein; 585 AA.
XX
AC AAR05318;
XX
DT 08-OCT-1990 (first entry)
XX
DE Human serum albumin gene product.
XX
KW Human serum albumin; HSA-A; yeast; ds.
XX
OS Homo sapiens.
XX
PN JP02117384-A.
XX
AC AAR08457;

```

RESULT 3

AAR08457

ID AAR08457 standard; Protein; 585 AA.

XX

AC AAR08457;

XX 25-MAR-2003 (updated)
 DT 16-APR-1991 (first entry)
 XX Human serum albumin.
 XX HSA; foinding; ss.
 XX Homo sapiens.
 OS Key Location/Qualifiers
 FH 123..303 /label= A
 FT Region
 FT 1..303 /label= B
 FT 123..585 /label= C
 FT
 XX JP02227079-A.
 PN 10-SEP-1990.
 XX 25-AUG-1989; 89JP-0217540.
 XX 25-AUG-1989; 89JP-0217540.
 XX (TOFU) TONEN CORP.
 PA
 XX WPI; 1990-317325/42.
 DR N-PSDB; AAQ06099.
 DR
 XX New human serum albumin fragments - used to bond medicines and for
 PT stable folding of protein(s).
 XX
 PS Claim 1; Fig 8; 24pp; Japanese.
 XX
 CC Fragments A-C of HSA are expressed as fusion proteins with the
 CC signal peptide of E. coli alkaline phosphatase. The fragments are
 CC selected for their specific properties. The C-terminal truncated
 CC fragment, B, does not bind long-chain fatty acids but does bind to
 CC various medicines at the central region. The N-terminal truncated
 CC fragment, C, has good stability in protein folding. The central
 CC segment, A, has characteristics of both B and C.
 CC See also AAQ06096-Q06098.
 CC (Updated on 25-MAR-2003 to correct PD field.)
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC
 XX Sequence 585 AA;
 SQ
 Query Match 100.0%; Score 3103; DB 11; Length 585;
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKSEVAHRFKDLGSENFKALVLIAPAOYLQOCFEDHVKLVNEVTFPAKTCVADESAAE 60
 DB 1 DAHKSEVAHRFKDLGSENFKALVLIAPAOYLQOCFEDHVKLVNEVTFPAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKORPERNECFLOHKDNPPLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKORPERNECFLOHKDNPPLRLVRPEV 120
 QY 121 DVNCTAFHNEETFLKKLYLVEIARRHPYFAPPELLFAKRYKAAFTCCQAAADKAACLLP 180
 DB 121 DVNCTAFHNEETFLKKLYLVEIARRHPYFAPPELLFAKRYKAAFTCCQAAADKAACLLP 180
 QY 181 KLDELDRDEGKASSAKORLKCASLQKFCGERAFKAWAVARLSORFPFAEFAEVSCLVTDITK 240
 DB 181 KLDELDRDEGKASSAKORLKCASLQKFCGERAFKAWAVARLSORFPFAEFAEVSCLVTDITK 240
 QY 241 VHTCECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVDENMPA 300
 DB 241 VHTCECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVDENMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPOYSVVLRLAKTYETTLLEKC 360
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPOYSVVLRLAKTYETTLLEKC 360
 QY 361 CAADPHECYAKVFDEPKPLVEEPONLIKONCELFEOQLGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADPHECYAKVFDEPKPLVEEPONLIKONCELFEOQLGEYKFNALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRLNLGKVGSKCKCKHPKAKRMPKCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRLNLGKVGSKCKCKHPKAKRMPKCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDETYVPKFNATETTFHADICTLSEKEROIKKOTALLVELVKHKPKAT 540
 DB 481 LVNRRPCFSALEVDETYVPKFNATETTFHADICTLSEKEROIKKOTALLVELVKHKPKAT 540
 QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
 DB 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
 RESULT 4
 AAR80301
 ID AAR80301 standard; Protein; 585 AA.
 XX
 AC AAR80301;
 XX
 DT 25-MAR-2003 (updated)
 DT 17-JAN-1996 (first entry)
 XX
 XX Human serum albumin.
 XX
 XX Serum albumin; HSA; aspartyl protease-3; Yap3p;
 XX Saccharomyces cerevisiae.
 XX
 OS Homo sapiens.
 XX
 PN KO9523857-A1.
 XX
 PD 08-SEP-1995.
 XX
 PF 01-MAR-1995; 95WO-GB00434.
 PR 05-MAR-1994; 94GB-0004270.
 XX
 PA (DELZ) DELTA BIOTECHNOLOGY LTD.
 XX
 PI Kerrywilliams SM, Gilbert SC;
 XX
 XX WPI; 1995-320572/41.
 DR N-PSDB; AAQ98695.
 XX
 PT Yeast with reduced levels of aspartyl protease 3 proteolytic
 PT activity - used to secrete human albumin without prodn. Of the 45
 PT KD fragment
 XX
 PS Example 1; Page 26-28; 50pp; English.
 XX
 CC The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected
 CC to site-directed mutagenesis to investigate the role of
 CC endoproteases in the generation of a 45 kDa albumin fragment obtd.
 CC when the cDNA is expressed in S. cerevisiae. Mutations were: R410A;
 CC L407A, L408V, V409A; and R410A, K413Q, K414Q. The latter set of
 CC mutations, especially, improved stability of HSA to yeast Yap3p
 CC proteolytic cleavage, allowing increased prodn. of recombinant HSA.
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 585 AA;
 Query Match 100.0%; Score 3103; DB 16; Length 585;
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGKGFENKALVLIAPAOYLQCCPDEHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHRFKDLGKGFENKALVLIAPAOYLQCCPDEHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
 QY 121 DVMTAFHDNEETFLKXLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 DB 121 DVMTAFHDNEETFLKXLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 QY 181 KDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEFVSKLVTDLT 240
 DB 181 KDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEFVSKLVTDLT 240
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVWLLRLAKTYETTLK 360
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVWLLRLAKTYETTLK 360
 QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLIKNCCELFEQGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADPHCEYAKVDFEFPKPLVEEPQNLIKNCCELFEQGEYKFNALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRNLGKVGSKCKGHPKAPMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRNLGKVGSKCKGHPKAPMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540
 DB 481 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540
 QY 541 KEQLKAVMDDFAAAFVEKCKCCKADDKETCFABEGKKLVAAASQAALGL 585
 DB 541 KEQLKAVMDDFAAAFVEKCKCCKADDKETCFABEGKKLVAAASQAALGL 585

RESULT 5

AAO20111

ID AAO20111 standard; Protein; 585 AA.

XX AC AAO20111;

XX DT 06-AUG-2002 (first entry)

XX DE HSA protein sequence related to the growth hormone protein.

XX KW Serum albumin-growth hormone fusion protein; growth hormone;

XX KW Down's syndrome.

XX OS Unidentified.

XX PN KR99076789-A.

XX PD 15-OCT-1999.

XX PF 25-JUN-1998; 98KR-0704914.

XX PR 30-DEC-1995; 95GB-0026733.

XX PR 19-DEC-1996; 96WO-GB03164.

XX PA (DELZ) DELTA BIOTECHNOLOGY LTD.

XX XX WPI; 1997-363680/55.

XX XX N-PSDB; AAK99568.

XX PT Serum albumin-growth hormone fusion protein - useful to treat growth

XX FT hormone related diseases, e.g. Down's syndrome

XX PS Disclosure; Fig 6; 2lpp; Korean.
 XX CC The invention relates to a serum albumin-growth hormone fusion protein -
 CC useful to treat growth hormone related diseases such as Down's syndrome.
 CC This sequence represents a HSA protein related to the serum albumin-
 CC growth hormone protein of the invention.
 XX SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 18; Length 585;
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGKGFENKALVLIAPAOYLQCCPDEHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHRFKDLGKGFENKALVLIAPAOYLQCCPDEHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
 QY 121 DVMTAFHDNEETFLKXLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 DB 121 DVMTAFHDNEETFLKXLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 QY 181 KDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEFVSKLVTDLT 240
 DB 181 KDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEFVSKLVTDLT 240
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVWLLRLAKTYETTLK 360
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVWLLRLAKTYETTLK 360
 QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLIKNCCELFEQGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADPHCEYAKVDFEFPKPLVEEPQNLIKNCCELFEQGEYKFNALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRNLGKVGSKCKGHPKAPMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRNLGKVGSKCKGHPKAPMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540
 DB 481 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540
 QY 541 KEQLKAVMDDFAAAFVEKCKCCKADDKETCFABEGKKLVAAASQAALGL 585
 DB 541 KEQLKAVMDDFAAAFVEKCKCCKADDKETCFABEGKKLVAAASQAALGL 585

RESULT 6

AA84873

ID AA84873 standard; protein; 585 AA.

XX AC AA84873;

XX DT 08-AUG-2000 (first entry)

XX DE Amino acid sequence of a human albumin protein.

XX KW Human; albumin; ischemic state; serum protein; metal ion salt;

XX KW perioperative ischemia; ischemia; myocardial infarction;

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Modified-site 1

/note= "optionally acetylated, and claimed under claim 56"

WO200020840-A1.

13-APR-2000.

01-OCT-1999; 99WO-JS22905.

02-OCT-1998; 98US-0102738.

02-OCT-1998; 98US-0165926;

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

.....

detecting and quanti

Disclosure; Page 97-100; 105pp; English.

The present sequence represents human albumin protein. The specification describes a method for the continuous detection of ischemic states. The method comprises detecting and quantifying the existence of an alteration of the serum protein albumin. The method comprises contacting a biological sample containing albumin from the patient with an excess quantity of a metal ion salt, where the metal ion binds to the N-terminus of naturally occurring human albumin, to form a mixture containing bound metal ions and unbound metal ions, and then determining the amount of metal ions bound to the albumin N-terminus. The amount of bound metal ions is correlated to a known value to determine the occurrence or non-occurrence of an ischemic event. The methods are useful for detection of ischemic states. The methods are also useful for distinguishing perioperative ischemia from ischemia caused by, amongst other things, myocardial infarctions and progressive coronary artery disease.

Sequence	585 AA;
SQ	

```
Query Match 100.0%; Score 3103; DB 21; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1 DAHKSEVAHRFKDOLGEENFKALVLI AFAOYLQQCPFFEDHVKL VNEVTSEAKTVADESAE 60
Db	1 DAHKSEVAHRFKDOLGEENFKALVLI AFAOYLQQCPFFEDHVKL VNEVTSEAKTVADESAE 60

Qy	61	NCDSLH	LT	FGDK	CL	TV	AT	LT	RET	VG	EM	DC	CA	KQ	EP	NE	CF	LQ	KD	NN	PL	VR	PE	VE	120
Db	61	NCDSLH	LT	FGDK	CL	TV	AT	LT	RET <td>VG <td>EM <td>DC <td>CA <td>KQ <td>EP <td>NE <td>CF <td>LQ <td>KD <td>NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	VG <td>EM <td>DC <td>CA <td>KQ <td>EP <td>NE <td>CF <td>LQ <td>KD <td>NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td>	EM <td>DC <td>CA <td>KQ <td>EP <td>NE <td>CF <td>LQ <td>KD <td>NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td></td></td></td></td></td></td></td></td></td>	DC <td>CA <td>KQ <td>EP <td>NE <td>CF <td>LQ <td>KD <td>NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td></td></td></td></td></td></td></td></td>	CA <td>KQ <td>EP <td>NE <td>CF <td>LQ <td>KD <td>NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td></td></td></td></td></td></td></td>	KQ <td>EP <td>NE <td>CF <td>LQ <td>KD <td>NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td></td></td></td></td></td></td>	EP <td>NE <td>CF <td>LQ <td>KD <td>NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td></td></td></td></td></td>	NE <td>CF <td>LQ <td>KD <td>NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td></td></td></td></td>	CF <td>LQ <td>KD <td>NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td></td></td></td>	LQ <td>KD <td>NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td></td></td>	KD <td>NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td></td>	NN <td>PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td></td>	PL <td>VR <td>PE <td>VE <td>120</td> </td></td></td>	VR <td>PE <td>VE <td>120</td> </td></td>	PE <td>VE <td>120</td> </td>	VE <td>120</td>	120

	Qy	Db
121	DVMCTAFHDNEETFLKKLYEIAIRRHPPYAPPELLFFAKRYKAATTECCQADKAACLIP 180	DVMCTAFHDNEETFLKKLYEIAIRRHPPYAPPELLFFAKRYKAATTECCQADKAACLIP 180

Qy	181	KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARJSQRPFAEPAEVSKLVTDLTK	240
Db	181	KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARJSQRPFAEPAEVSKLVTDLTK	240

Qy	241	VHTECCHGOLLECA	DDRADLAKYICENQDSISSKLUKECEKPLEKSHCIAE	300
		=====	=====	
Db	241	VHTECCHGOLLECA	DDRADLAKYICENQDSISSKLUKECEKPLEKSHCIAE	300
		=====	=====	

[illegible]

361 CAADPHCEYAKVEDEFKDLVEEPQNLIKQNCSELFQJGEYKFQNALLYRYTKKVPQVST 420

361	CAAADPHECYAKVDFE	KPLVEEPQNI	KONCELF	EQLG	EYKIQ	NALLV	RYTKK	PQVST	420
421	PTLVSVSNLGVGSKCK	HPKAPKMP	CAEDYLSW	NQLCVL	HEKTPV	SDRYTK	CKCTS		480
421	PTLVSVSNLGVGSKCK	HPKAPKMP	CAEDYLSW	NQLCVL	HEKTPV	SDRYTK	CKCTS		480
481	LVNRRPCFSALVEDET	YVPKEFNA	ETTFH	ACICTL	SEKERO	KKQTAL	VELVKKH	KPKAT	540
481	LVNRRPCFSALVEDET	YVPKEFNA	ETTFH	ACICTL	SEKERO	KKQTAL	VELVKKH	KPKAT	540
541	KQOLKAVMDDFAAF	VEKCKCK	ADDKETCF	AE	BEGK	KLVAAS	QAALGL	585	
541	KQOLKAVMDDFAAF	VEKCKCK	ADDKETCF	AE	BEGK	KLVAAS	QAALGL	585	

RESULT 7
AAV83946
ID AAV83946 standard: Protein: 585 AA.

AC AAY83946;
 XX
 DT 28-JUL-2000 (first entry)

DE Yeast codon-biased recombinant human serum albumin protein.
XX
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
KW overlapping oligonucleotide; expression vector.

OS	Homo sapiens.
OS	Synthetic.
XX	

XX
XX
PD
XX

XX
PR 17-JUN-1998; 98CN-0102506.
XX

XX
PI
Li S, Lu D;
XX

DR N-PSDB; AAA10091.
XX
PT Process for preparing recombined human serum albumin - which comprises

XX
PS Disclosure; Fig 1; 44pp; Chinese.
XX

CC albumin (HSA) in yeast by altering the coding sequence of HSA to
CC comprise a yeast codon bias. The complete HSA gene (AA10091) was
CC generated as three synthetic fragments (AA10092-AA10094) joined by

CC overlapping oligonucleotide fragments that were extended. This sequence
CC represents the complete sequence of the HSA encoded by the human gene
CC with a yeast codon bias. The invention also covers a recombinant

CC vector and the process for producing human serum albumin in the yeast
CC host cell, especially in secretory mode.

```
Query Match      100.0%; Score 3103; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
```

QY 1 DAHKSEVAHRFKDILGEENFKALVJIAFAQYLOQCPFDHVKLVNEVTEFAKTCVADESAE 60

Qy	1	DAHKSEVAHRFKDGLGENFKALVLIAPAOYLQCCPEDHVKLVNVEYFAKTCVADESAS	60
Db	1	DAHKSEVAHRFKDGLGENFKALVLIAPAOYLQCCPEDHVKLVNVEYFAKTCVADESAS	60
Qy	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDKPNLPRJVRPEV	120
Db	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDKPNLPRJVRPEV	120
Qy	121	DVMCTAFHDNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP	180
Db	121	DVMCTAFHDNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP	180
Qy	181	KLDLDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQSPKAEPAEVSCLVTD-TK	240
Db	181	KLDLDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQSPKAEPAEVSCLVTD-TK	240
Qy	241	VHTECCGDLLECCADRADLAKYICENQDSISSKLKECCCEKPLLEKSHCIAEVENDENMPA	300
Db	241	VHTECCGDLLECCADRADLAKYICENQDSISSKLKECCCEKPLLEKSHCIAEVENDENMPA	300
Qy	301	DLPSLAADFVESKDVCKNYABAKDVFGLMFLVEYARRHPDYSVLLLRLLAKTYETTLKRC	360
Db	301	DLPSLAADFVESKDVCKNYABAKDVFGLMFLVEYARRHPDYSVLLLRLLAKTYETTLKRC	360
Qy	361	CAAADPHECVAKVDFDEPKPLVEEPQNLIKONCELFEOLGEYKFNALLVRYTKKVPQVST	420
Db	361	CAAADPHECVAKVDFDEPKPLVEEPQNLIKONCELFEOLGEYKFNALLVRYTKKVPQVST	420
Qy	421	PTLVEVSRNLGKVGSKCKHPEAKRMPKPCADYLSVVLNQLCVLHEKTPVSDRVTKCCTES	480
Db	421	PTLVEVSRNLGKVGSKCKHPEAKRMPKPCADYLSVVLNQLCVLHEKTPVSDRVTKCCTES	480
Qy	481	LVNRRPFCFSALEVDVTVPKFNAETFTPHADICTLSEKERQIKKQTAUVLVKHKPKAT	540
Db	481	LVNRRPFCFSALEVDVTVPKFNAETFTPHADICTLSEKERQIKKQTAUVLVKHKPKAT	540
Qy	541	KEQLKAVMDDFAAVFEKCKCKADDKETCFABEGKKLVAASQAALGL	585
Db	541	KEQLKAVMDDFAAVFEKCKCKADDKETCFABEGKKLVAASQAALGL	585
RESULT 9			
AAE13399			
ID	AAE13399	standard; Protein; 585 AA.	
XX	AAE13399;		
AC	AAE13399;		
XX	12-FEB-2002	(first entry)	
DE	Human albumin (HA) protein.		
XX	Human; albumin; HA; fusion protein; immune system disorder; syphilis;		
KW	transplant rejection; blood related disorder; myocardial infarction;		
KW	hyperproliferative disorder; acute myeloid leukaemia; renal disorder;		
KW	glomerulonephritis; cardiovascular disease; arrhythmia; rhinitis;		
KW	respiratory disorder; neurological disease; Alzheimer's disease;		
KW	endocrine disorder; pheochromocytoma; reproductive system disorder;		
KW	measles; gastrointestinal disorder; irritable bowel syndrome; HIV;		
KW	human immunodeficiency virus; wound healing; renal cell carcinoma;		
KW	melanoma; gene therapy.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	Domain	54..61	
FT	Domain	/label= Loop_I	
FT	Domain	76..89	
FT	Domain	/label= Loop_II	
FT	Domain	92..100	
FT	Domain	/label= Loop_III	
FT	Domain	170..176	
FT	Domain	/label= Loop_IV	
FT	Domain	247..252	

FT	Domain	/label= Loop_V	
FT	Domain	266..277	
FT	Domain	/label= Loop_VI	
FT	Domain	280..288	
FT	Domain	/label= Loop_VII	
FT	Domain	362..368	
FT	Domain	/label= Loop_VIII	
FT	Domain	439..447	
FT	Domain	/label= Loop_IX	
FT	Domain	461..475	
FT	Domain	/label= Loop_X	
FT	Domain	478..486	
FT	Domain	/label= Loop_XI	
FT	Domain	560..566	
FT	Domain	/label= Loop_XII	
XX	WO200179258-A1.		
PN	25-OCT-2001.		
PD	12-APR-2001; 2001WO-US12008.		
ZF	12-APR-2000; 2000US-229358P.		
XX	25-APR-2000; 2000US-199384P.		
PR	21-DEC-2000; 2000US-256931P.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
PA	(PRIN-) PRINCIPIA PHARM CORP.		
XX	Rosen CA, Sadeghi H, Prior CP, Turner AJ;		
PI	WPI; 2001-602931/68.		
XX	N-PSDB; AAD22287.		
DR	Albumin fusion proteins comprising a therapeutic protein and albumin,		
XX	useful in the treating metastatic renal cell carcinoma, metastatic		
PT	melanoma, malignant melanoma, renal cell carcinoma, HIV (human		
PT	immunodeficiency virus) or infection -		
XX	Claim 1; Fig 9; 325pp; English.		
PS	The invention relates to albumin fusion proteins comprising therapeutic		
XX	protein and human albumin (HA). The albumin fusion proteins are useful		
CC	in the treatment, prevention, diagnosis, and/or detection of diseases,		
CC	disorders such as immune system disorders (transplant rejection); blood		
CC	related disorders (myocardial infarction); hyperproliferative disorders		
CC	(childhood acute myeloid leukaemia); renal disorder (glomerulonephritis);		
CC	cardiovascular disorders (arrhythmias); respiratory disorders		
CC	(non-allergic rhinitis); neurological diseases (Alzheimer's disease);		
CC	endocrine disorders (pheochromocytoma); reproductive system disorders		
CC	(syphilis); infectious diseases (measles); gastrointestinal disorders		
CC	(irritable bowel syndrome) and wound healing. The albumin fusion		
CC	proteins are also used in the treatment of metastatic renal cell		
CC	carcinoma, metastatic melanoma, malignant melanoma and HIV (human		
CC	immunodeficiency virus) infection. Nucleic acid encoding albumin fusion		
CC	protein is useful in gene therapy. The present sequence is human		
CC	albumin (HA) protein.		
XX	Sequence 585 AA;		
SQ	Query Match 100.0%; Score 3103; DB 22; Length 585;		
	Best Local Similarity 100.0%; Pred. No. 9,5e-255;		
	Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	DAHKSEVAHRFKDGLGENFKALVLIAPAOYLQCCPEDHVKLVNVEYFAKTCVADESAS	60
Db	1	DAHKSEVAHRFKDGLGENFKALVLIAPAOYLQCCPEDHVKLVNVEYFAKTCVADESAS	60
Qy	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDKPNLPRJVRPEV	120
Db	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDKPNLPRJVRPEV	120
Qy	121	DVMCTAFHDNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP	180

```

Db      |||||
121  DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQADKAACLLP 180
QY      |||||
181  KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSCLVTDITK 240
Db      |||||
181  KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSCLVTDITK 240
QY      |||||
241  VHTCCCHGDLLECCADRADLAKYICENODSISKKLKECCCKPPLLEKSHCIAEVENDEMPA 300
Db      |||||
241  VHTCCCHGDLLECCADRADLAKYICENQDSISKKLKECCCKPPLLEKSHCIAEVENDEMPA 300
QY      |||||
301  DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTTLEKC 360
Db      |||||
301  DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTTLEKC 360
QY      |||||
361  CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNQALVRYTKKVPQVST 420
Db      |||||
361  CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNQALVRYTKKVPQVST 420
QY      |||||
421  PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db      |||||
421  PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY      |||||
481  LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540
Db      |||||
481  LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540
QY      |||||
541  KEOLKAVMDDFAAVFEKCKCKADDDKTCFAEFGKKLVAASQAALGL 585
Db      |||||
541  KEOLKAVMDDFAAVFEKCKCKADDDKTCFAEFGKKLVAASQAALGL 585

```

RESULT 20

AA052567

ID AA052567 standard; Protein: 585 AA.

XX

AC AA052567;

XX

DT 05-FEB-2002 (first entry)

XX

DE Mature human serum albumin.

XX

Human; serum albumin; HA; antiinflammatory; immunosuppressive; cardiac;
nootropic; neuroprotective; gene therapy; immune disorder; wound healing;
hyperproliferative disorder; renal disorder; cardiovascular disorder;
respiratory disorder; neurological disease; endocrine disorder;
reproductive system disorder; infectious disease;
gastrointestinal disorder.

XX

OS Homo sapiens.

XX

PN W0200179444-A2.

XX

PD 25-OCT-2001.

XX

XX 12-APR-2001; 2001WO-US12013.

PF

PR 12-APR-2000; 2000US-229358P.

PR

PR 25-APR-2000; 2000US-199384P.

PR

XX 21-DEC-2000; 2000US-256931P.

XX

(HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Haseltine WA;

XX

DR WPI; 2001-616755/71.

DR

XX N-PSDB; ABA03057.

XX

Albumin fusion proteins comprising a therapeutic protein and albumin,
useful in the treating immune system disorders (e.g. transplant
rejection), blood related disorders (e.g. myocardial infarction) and
hyperproliferative disorders -

PT

PT

XX

DE Human albumin (HA).

Claim 1; Fig 15; 606pp; English.

XX The present invention relates to albumin fusion proteins, which comprise
a therapeutic protein and albumin. The present sequence is the protein
sequence for mature human serum albumin (HA), which was used to generate
the fusion proteins of the present invention. The albumin fusion proteins
are useful in the treatment, prevention, diagnosis, and/or detection of
diseases/disorders such as immune system disorders (e.g. transplant
rejection), blood related disorders (e.g. myocardial infarction),
hyperproliferative disorders (e.g. childhood acute myeloid leukemia),
renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.
arrhythmias), respiratory disorders (e.g. non-allergic rhinitis),
neurological diseases (e.g. Alzheimer's disease), endocrine disorders
(e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis),
infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
irritable bowel syndrome) and wound healing.

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 22; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHREFKDLGEENFKALVLIAPQYVLCQCFEDHVKLVNEVTEFAKTCVADESAAE 60

Db 1 DAHKSEVAHREFKDLGEENFKALVLIAPQYVLCQCFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDNPNLPVLVRPEV 120

Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDNPNLPVLVRPEV 120

QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQADKAACLLP 180

Db 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQADKAACLLP 180

QY 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSCLVTDITK 240

Db 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSCLVTDITK 240

QY 241 VHTCCCHGDLLECCADRADLAKYICENODSISKKLKECCCKPPLLEKSHCIAEVENDEMPA 300

Db 241 VHTCCCHGDLLECCADRADLAKYICENODSISKKLKECCCKPPLLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTTLEKC 360

Db 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTTLEKC 360

QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNQALVRYTKKVPQVST 420

Db 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNQALVRYTKKVPQVST 420

QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Db 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540

Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540

QY 541 KEOLKAVMDDFAAVFEKCKCKADDDKTCFAEFGKKLVAASQAALGL 585

Db 541 KEOLKAVMDDFAAVFEKCKCKADDDKTCFAEFGKKLVAASQAALGL 585

RESULT 11

AA0513129

ID AA0513129 standard; Protein: 585 AA.

XX

AC AA0513129;

XX

DT 28-JAN-2002 (first entry)

XX

DE Human albumin (HA).

XX Human; albumin; HA; fusion protein; therapeutic protein; vulnereary;
KW immune system disorder; transplant rejection; blood related disorder;
KW myocardial infarction; hyperproliferative disorder; glomerulonephritis;
KW childhood acute myeloid leukemia; cardiovascular disorder; arrhythmia;
KW respiratory disorder; gene therapy; non-allergic rhinitis; neurotropic;
KW neurological disease; Alzheimer's disease; reproductive system disorder;
KW endocrine disorder; pheochromocytoma; infectious disease; antiarthritic;
KW measles; gastrointestinal disorder; irritable bowel syndrome; syphilis;
KW wound healing; antiinflammatory; immunosuppressive; neuroprotective;
KW cardiac; cytostatic; antileukemic; antineumatic; antimicrobial;
KW renal disorder.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Domain 54..61 /label= Loop_I
FT Domain 76..89 /label= Loop_II
FT Domain 92..100 /label= Loop_III
FT Domain 170..176 /label= Loop_IV
FT Domain 247..252 /label= Loop_V
FT Domain 266..277 /label= Loop_VI
FT Domain 280..288 /label= Loop_VII
FT Domain 362..368 /label= Loop_VIII
FT Domain 439..447 /label= Loop_IX
FT Domain 461..475 /label= Loop_X
FT Domain 478..486 /label= Loop_XI
FT Domain 560..566 /label= Loop_XII
PN WO200179443-A2.
XX 25-OCT-2001.
XX 12-APR-2001; 2001WO-US11:924.
XX 12-APR-2000; 2000US-229358P.
XX 25-APR-2000; 2000US-199384P.
XX 21-DEC-2000; 2000US-256931P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Haseltine WA;
XX WPI; 2001-616754/71.
XX N-PSDB; AAD21638.
XX Albumin fusion proteins comprising a therapeutic protein and albumin,
PT useful in the treating immune system disorders (e.g. transplant
PT rejection), blood related disorders (e.g. myocardial infarction) and
PT hyperproliferative disorders -
XX
XX Claim 1; Fig 9; 380pp; English.
XX The invention relates to albumin fusion proteins comprising therapeutic
CC protein and human albumin (HA). Therapeutic protein fused to albumin
CC have an extended shelf-life. The albumin fusion proteins are useful in
CC the treatment, prevention, diagnosis and/or detection of diseases,
CC disorders such as immune system disorders (e.g. transplant rejection),
CC blood related disorders (e.g. myocardial infarction), hyperproliferative
CC disorders (e.g. childhood acute myeloid leukaemia), renal disorders
CC (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),

CC respiratory disorders (e.g. non-allergic rhinitis), neurological
CC diseases (e.g. Alzheimer's disease), endocrine disorders (e.g.
CC pheochromocytoma), reproductive system disorders (e.g. syphilis),
CC infectious diseases (e.g. measles), gastrointestinal disorders, (e.g.
CC irritable bowel syndrome) and wound healing. Nucleic acids encoding
CC albumin fusion protein is used in gene therapy. The present sequence
CC is human albumin (HA) protein.
XX
SQ Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 22; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFDLGGDKLVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 1 DAHKSEVAHRFDLGGDKLVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
QY 61 NCDKSLHTLFGDKLVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
QY 121 DVMTAFHNEETFLKKYLYEIAIRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMTAFHNEETFLKKYLYEIAIRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECADDRADLAKYICENQDSISSKLECCCKPPLLEKSHGICIAEVNDEMPA 300
DB 241 VHTCCCHGDLLECADDRADLAKYICENQDSISSKLECCCKPPLLEKSHGICIAEVNDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYEARHPDYSVLLLRLLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYEARHPDYSVLLLRLLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEPKPLVEBPQNLIKONCELPQOLGEYKFQNALIVRYTKKVPQVST 420
DB 361 CAAADPHECYAKVDFEPKPLVEBPQNLIKONCELPQOLGEYKFQNALIVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480
DB 421 PTLVEVSRLNGKVGSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480
QY 481 LVNRRFCFSALEVDYTPKPFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 540
DB 481 LVNRRFCFSALEVDYTPKPFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 540
QY 541 KEOLKAVMDDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
DB 541 KEOLKAVMDDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
RESULT 12
AAE12403
ID AAE12403 standard; Protein; 585 AA.
XX
AC AAE12403;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human albumin (HA).
XX Human; albumin; HA; immune system disorder; transplant rejection;
KW blood related disorder; myocardial infarction; glomerulonephritis;
KW hyperproliferative disorder; childhood acute myeloid leukaemia;
KW renal cell carcinoma; cardiovascular disorder; vulnereary; melanoma;
KW arrhythmia; respiratory disorder; non-allergic rhinitis; antileukemic;
KW neurological disease; Alzheimer's disease; endocrine disorder; measles;
KW pheochromocytoma; reproductive system disorder; neuroprotective; syphilis;
KW infectious disease; gastrointestinal disorder; wound healing; neurotropic;

KW irritable bowel syndrome; HIV; human immunodeficiency virus infection;
KW cytostatic; antiinflammatory; gene therapy; immunosuppressive; cardiant;
KW antiarthritic; antirheumatic; renal disorder; antimicrobial.

OS Homo sapiens.

XX Key Location/Qualifiers
PH Domain 54..61
FT /label= Loop_I
FT Domain 76..89
FT /label= Loop_II
FT Domain 92..100
FT /label= Loop_III
FT Domain 170..176
FT /label= Loop_IV
FT Domain 247..252
FT /label= Loop_V
FT Domain 266..277
FT /label= Loop_VI
FT Domain 280..288
FT /label= Loop_VII
FT Domain 362..368
FT /label= Loop_VIII
FT Domain 439..447
FT /label= Loop_IX
FT Domain 461..475
FT /label= Loop_X
FT Domain 478..486
FT /label= Loop_XI
FT Domain 560..566
FT /label= Loop_XII

WO200179480-A1.

25-OCT-2001.

12-APR-2001; 2001WO-US11991.

12-APR-2000; 2000US-229358P.

25-APR-2000; 2000US-199384P.

21-DEC-2000; 2000US-256931P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Haseltine WA;

WPI; 2001-616756/71.

N-PSDB; AAD20005.

Albumin fusion proteins comprising a therapeutic protein and albumin,
useful in the treating metastatic renal cell carcinoma, metastatic
melanoma, malignant melanoma, renal cell carcinoma, HIV (human
immunodeficiency virus) or infection -

Claim 1; Fig 9; 394pp; English.

The invention relates to human albumin (HA) fusion proteins and their
corresponding nucleic acid sequences. Therapeutic proteins fused to
albumin or its fragments have an extended shelf-life. The albumin
fusion proteins are useful in the treatment, prevention, diagnosis,
and/or detection of diseases, disorders such as immune system
disorders (e.g. transplant rejection), blood related disorders (e.g.
myocardial infarction), hyperproliferative disorders (e.g. childhood
acute myeloid leukaemia, metastatic renal cell carcinoma, metastatic
melanoma, malignant melanoma, renal cell carcinoma), renal disorders
(e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),
respiratory disorders (e.g. non-allergic rhinitis), neurological
diseases (e.g. Alzheimer's disease), endocrine disorders (e.g.
pheochromocytoma), reproductive system disorders (e.g. syphilis),
infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
irritable bowel syndrome), HIV (human immunodeficiency virus) infection
and wound healing. Nucleic acids encoding albumin fusion protein is
used in gene therapy. The present sequence is human albumin.

XX Sequence 585 AA;
SQ Query Match 100.0%; Score 3103; DB 22; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255; Mismatches 0; Indels 0; Gaps 0;
Matches 585; Conservative 0;
QY 1 DAHKSEVAHREFKDLGEENFKALVLIIFAQYLQOCFFEDHVKLVNVEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHREFKDLGEENFKALVLIIFAQYLQOCFFEDHVKLVNVEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOPEPERNECFLOHKDNDPNLPLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOPEPERNECFLOHKDNDPNLPLVRPEV 120
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180
DB 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPAKFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPAKFAEVSCLVTDLT 240
QY 241 VHTCCGHDLLLECADDADLAKYICENODSISLKECCCKPILLESKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLLECADDADLAKYICENODSISLKECCCKPILLESKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360
QY 361 CAADPHECYAKVDFEPKPLVEEPONLIKQNCLEPEQGEYKFNALLVRYTKVPQVST 420
DB 361 CAADPHECYAKVDFEPKPLVEEPONLIKQNCLEPEQGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSRVTKCCTES 480
QY 481 LVNRRPCFSALEVDDETYVPKFNAEFTTFHADICTLSSEKQIKQKQALVELVHKPKAT 540
DB 481 LVNRRPCFSALEVDDETYVPKFNAEFTTFHADICTLSSEKQIKQKQALVELVHKPKAT 540
QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKETCFPAEEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDDFAAFVEKCKKADDDKETCFPAEEGKKLVAASQAALGL 585
RESULT 13
AAE08578
ID AAE08578 standard; Protein; 585 AA.
XX
AC AAE08578;
XX
DT 19-NOV-2001 (first entry)
XX
DE Human serum albumin (HSA).
XX
KW Human; albumin; cancer; cell proliferation; drug screening; biopsy.
XX
OS Homo sapiens.
XX
PN US6274305-B1.
XX
PD 14-AUG-2001.
XX
PF 19-DEC-1996; 96US-0769746.
XX
PR 19-DEC-1996; 96US-0769746.
XX
PA (TUFT) UNIV TUFTS.
XX
PI Sonnenschein C, Soto AM;

XX WPI; 2001-540371/60.
 DR N-PSDB; AAD11488.
 XX
 PT Measuring human cell proliferation, useful in drug screening to
 PT determine the potential for inhibiting cancer cell proliferation and
 PT for evaluating biopsied tumors, comprises employing albumin-derived
 PT peptide -
 XX
 XX
 PS Claim 1; Fig 1; 20pp; English.
 XX
 XX The invention related to a method for testing cancer cells. The method is
 CC useful for measuring human cancer cell proliferation, particularly for
 CC determining the potential for inhibiting cancer cells proliferation using
 CC albumin-derived peptides. The invention is also useful for drug screening
 CC assays, as well as for evaluating biopsied tumors. The present sequence
 CC is human serum albumin (HSA) related to the invention.
 XX
 XX
 SQ Sequence 585 AA;
 Query Match 100.0%; Score 3103; DB 22; Length 585;
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHREFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHREFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELFFAKRYKAAPTECCQAADKAACLLP 180
 DB 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELFFAKRYKAAPTECCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAPKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKQRLKCSLQKFGERAPKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
 QY 241 VHTCCGGDLLECADRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCGGDLLECADRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLK 360
 DB 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLK 360
 QY 361 CAADPHCEYAKVDFEKLVEEPQNLIKONCELFEOQGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADPHCEYAKVDFEKLVEEPQNLIKONCELFEOQGEYKFNALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDVETVYPKEFNABTFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
 DB 481 LVNRRPCFSALEVDVETVYPKEFNABTFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
 QY 541 KEQLKAVNMDPFAAFVEKCKCKADKDETCAFEKGKLVAAASQAALGL 585
 DB 541 KEQLKAVNMDPFAAFVEKCKCKADKDETCAFEKGKLVAAASQAALGL 585
 RESULT 14
 ID ABG71291
 XX ABG71291 standard; Protein; 585 AA.
 AC
 XX ABG71291;
 XX
 DT 08-JAN-2003 (first entry)
 XX

DE Glycosylated protein determination associated protein.
 XX Standard substance; accuracy control substance; glycosylated protein;
 KW glycosylated albumin; fructosamine; diabetes; antidiabetic.
 OS Unidentified.
 XX
 XX JP2002243731-A.
 PN 28-AUG-2002.
 XX
 XX 21-FEB-2001; 2001JP-0045085.
 PR 21-FEB-2001; 2001JP-0045085.
 PA {KOKJ-} KOKUSAI SHIYAKU KK.
 PA {YOSH } YOSHITOMI PHARM IND KK.
 XX WPI; 2002-744850/81.
 DR A standard substance for determination of glycosylated protein
 XX including glycosylated albumin and fructosamine, used in diagnosis of
 PT diabetes -
 PS Disclosure; Page 4; 6pp; Japanese.
 XX The present invention relates to a new standard and accuracy control
 CC substance for determination of glycosylated protein. The invention is
 CC useful for determination of glycosylated protein in the diagnosis of
 CC diabetes. Glycosylated albumin and fructosamine provide favourable
 CC dilution linearity. The present amino acid sequence represents the
 CC glycosylated protein determination associated protein as described in
 CC the invention.
 XX
 SQ Sequence 585 AA;
 Query Match 100.0%; Score 3103; DB 23; Length 585;
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHREFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHREFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELFFAKRYKAAPTECCQAADKAACLLP 180
 DB 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELFFAKRYKAAPTECCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAPKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKQRLKCSLQKFGERAPKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
 QY 241 VHTCCGGDLLECADRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCGGDLLECADRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLK 360
 DB 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLK 360
 QY 361 CAADPHCEYAKVDFEKLVEEPQNLIKONCELFEOQGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADPHCEYAKVDFEKLVEEPQNLIKONCELFEOQGEYKFNALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDVETVYPKEFNABTFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540

Db 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQALVELVKHKPKAT 540
541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALGL 585
541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALGL 585

RESULT 15
ABG63321
ID ABG63321 standard; protein: 585 AA.
XX AC ABG63321;
XX DT 27-AUG-2002 (first entry)
XX DE Human serum albumin (HSA) protein.
XX KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytosolic; antiinfertility; antiinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX OS Homo sapiens.
XX PS WO200177137-A1.
XX PD 18-OCT-2001.
XX PF 12-APR-2001; 2001WO-US11988.
XX PR 12-APR-2000; 2000US-229358P.
XX PR 25-APR-2000; 2000US-199384P.
XX PR 21-DEC-2000; 2000US-256931P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Haseltine WA;
XX DR WPI; 2002-010886/01.
XX DR N-PSDB; ABK93280.
XX PT New fusion protein for treating disease e.g. diabetes comprises an
XX PT albumin fused to a therapeutic protein -
XX PS Claim 1; Fig 15; 2102pp; English.
XX CC The present invention relates to albumin fusion proteins comprising a
XX CC therapeutic protein X and human albumin (HA), also known as human serum
XX CC albumin, HSA). The proteins are useful for treating a disease or
XX CC disorder that may be modulated by therapeutic protein X. The albumin
XX CC extends the shelf-life of protein X, and may increase its biological
XX CC in vitro/in vivo activity. The protein is useful for treating and
XX CC diagnosing disorders such as cancer, reproductive disorders, digestive
XX CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
XX CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
XX CC (e.g. diabetes), haematopoietic disorders, neural disorders
XX CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
XX CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
XX CC (e.g. osteoporosis, arthritis). The present sequence represents HSA
XX CC (HA) protein.
XX SQ Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 23; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGGENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKCVADESAE 60

Db 1 DAHKSEVAHRFKDLGGENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEYARRHPYFYAPELFFAKRYKAAFTTECCOAAKKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEYARRHPYFYAPELFFAKRYKAAFTTECCOAAKKAACLLP 180
QY 181 KLDELDECKASSAKORLKCASLQKGFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
Db 181 KLDELDECKASSAKORLKCASLQKGFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTTECHGDLLECADDDRADLAKYICENODSISKKLKECCEKPLEKSHCIAEVNDEMPA 300
Db 241 VHTTECHGDLLECADDDRADLAKYICENODSISKKLKECCEKPLEKSHCIAEVNDEMPA 300
QY 301 DLPSLAADFVESKDVCKYAEAKDVLGMFLYAYARRHPDYSVWJLLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKYAEAKDVLGMFLYAYARRHPDYSVWJLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFPKPLVBEOPNLKQNCLEPEQGEYKFNALLVRYTKVPQVST 420
Db 361 CAAADPHECYAKVDFEFPKPLVBEOPNLKQNCLEPEQGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPEAKMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQALVELVKHKPKAT 540
Db 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQALVELVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALGL 585

Search completed: October 27, 2003, 15:32:35
Job time : 87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:31:09 ; Search time 29 Seconds
(without alignments)
853.512 Million cell updates/sec

Title: US-09-832-929-18
Perfect score: 3103
Sequence: 1 DAHSEVAHRFKDLGSENFK.....TCFAEGKKLVAAQAALCL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	585	1	US-08-153-799-14
2	3103	100.0	585	2	US-08-702-572-2
3	3103	100.0	585	3	US-08-769-746-2
4	3103	100.0	610	2	US-08-797-689-2
5	3103	100.0	783	1	US-08-256-938-2
6	3103	100.0	787	1	US-08-256-938-4
7	3103	100.0	787	2	US-08-797-689-16
8	3099	99.9	609	1	US-08-222-619-3
9	3099	99.9	609	1	US-08-433-037-4
10	3099	99.9	609	4	US-08-897-956A-2
11	3099	99.9	609	5	PCT-US95-04075-3
12	3095	99.7	978	4	US-08-897-956A-3
13	3093	99.7	585	1	US-08-448-196A-3
14	3093	99.7	585	2	US-08-984-176-1
15	2458.5	79.2	583	1	US-08-448-196A-5
16	2450.5	79.0	583	1	US-08-448-196A-4
17	2432.5	78.4	583	1	US-08-448-196A-6
18	2426	78.2	584	1	US-08-448-196A-7
19	2389	77.0	582	1	US-08-134-638-1
20	1249.5	40.3	609	1	US-08-222-619-4
21	1249.5	40.3	609	5	PCT-US95-04075-4
22	1206.5	38.9	590	2	US-08-377-309-2
23	1206.5	38.9	590	3	US-09-186-723-2
24	1206.5	38.9	590	4	US-08-505-012-5
25	1206.5	38.9	590	4	US-09-186-949A-3
26	1206.5	38.9	590	4	US-08-758-757-2
27	1206.5	38.9	590	5	PCT-US96-00996-5

28	1206.5	38.9	609	4	US-09-186-949A-2	Sequence 2, Appl
29	1164.5	37.5	579	1	US-08-448-196A-8	Sequence 8, Appl
30	1055	34.0	599	1	US-08-222-619-2	Sequence 2, Appl
31	1055	34.0	599	3	US-08-221-767-24	Sequence 24, Appl
32	1055	34.0	599	5	PCT-US95-04075-2	Sequence 2, Appl
33	926	29.8	393	2	US-08-377-309-7	Sequence 7, Appl
34	926	29.8	393	3	US-09-186-723-7	Sequence 7, Appl
35	926	29.8	393	4	US-08-505-012-10	Sequence 10, Appl
36	926	29.8	393	4	US-09-186-949A-8	Sequence 8, Appl
37	926	29.8	393	4	US-08-758-757-7	Sequence 7, Appl
38	926	29.8	393	5	PCT-US96-00996-10	Sequence 10, Appl
39	777	25.0	324	4	US-08-505-012-12	Sequence 12, Appl
40	777	25.0	324	5	PCT-US96-00996-12	Sequence 12, Appl
41	777	25.0	325	2	US-08-377-309-8	Sequence 8, Appl
42	777	25.0	325	3	US-09-186-723-8	Sequence 8, Appl
43	777	25.0	325	4	US-08-505-012-11	Sequence 11, Appl
44	777	25.0	325	4	US-09-186-949A-9	Sequence 9, Appl
45	777	25.0	325	4	US-08-758-757-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-1310
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,572
FILING DATE: 11-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 95/23857
FILING DATE: 1-MAR-1995
APPLICATION NUMBER: GB 9404270.2
FILING DATE: 5-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Naomi Biewas
REGISTRATION NUMBER: 38,384
REFERENCE/DOCKET NUMBER: CE0114 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610/878/4294
TELEFAX: 610/878/4221
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-702-572-2

Query Match      100.0%; Score 3103; DB 2; Length 585;
Best Local Similarity 100.0%; Pred No. 9.3e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY    1 DAHKSEVAHRFKDGLGEENFKALVLIAPAQYLQOCPPEDHVKLVNEVFPAKTCVADSSAE   60
DB    1 DAHKSEVAHRFKDGLGEENFKALVLIAPAQYLQOCPPEDHVKLVNEVFPAKTCVADSSAE   60
QY    6: NCDKSHTLFGDKLCVTATILRETYGEMADCCAKOEPNERBECFLQHKKDDNPNLPRLVPPEV 120
DB    6: NCDKSHTLFGDKLCVTATILRETYGEMADCCAKOEPNERBECFLQHKKDDNPNLPRLVPPEV 120
QY    121 DVNCTAFHNEETFLKKLYIEIARRHPYPFYAPPELLFFAKRYKAFTGCCQAADKAACLLP 180
DB    121 DVNCTAFHNEETFLKKLYIEIARRHPYPFYAPPELLFFAKRYKAFTGCCQAADKAACLLP 180
QY    181 KLDELDEGKASSAKQRKLKCSLQKGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 240
DB    181 KLDELDEGKASSAKQRKLKCSLQKGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 240
QY    241 VHTCECHGDLLECADRDADLAKYICENQDISISKLKECKEKPLLEKSHCIAEVENDMPA 300
DB    241 VHTCECHGDLLECADRDADLAKYICENQDISISKLKECKEKPLLEKSHCIAEVENDMPA 300
QY    301 DLPSLAADPVESKDVCNKVAAEKDVFLGMFLYEYARRHPDYYSVVLLLRLLAKTYETTLEKC 360
DB    301 DLPSLAADPVESKDVCNKNTABAKDVFLGMFLYEYARRHPDYYSVVLLLRLLAKTYETTLEKC 360
QY    361 CAAADPHCYAKYVDFEFKPLVEEPQNLIKONCELFEQLGEYKFQNALLYRTTKVPQVST 420
DB    361 CAAADPHCYAKYVDFEFKPLVEEPQNLIKONCELFEQLGEYKFQNALLYRTTKVPQVST 420
QY    421 PTLVEYSRNLGVKGSKCKKHPEAKRWPCBAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB    421 PTLVEYSRNLGVKGSKCKKHPEAKRWPCBAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY    481 LVNRRFCFSALEYDETYYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
DB    481 LVNRRFCFSALEYDETYYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540

```

QY 541 KEQLKAVMDDDFAAFVEKCKCKADKCTCFABEGKLVAAASQAALGL 585
 |||||
 DB 541 KEQLKAVMDDDFAAFVEKCKCKADKCTCFABEGKLVAAASQAALGL 585
 |||||

RESULT 3
 US-08-769-746-2
 ; Sequence 2, Application US/08769746
 ; Patent No. 6274305
 ; GENERAL INFORMATION:
 ; APPLICANT: Sonnenschein, Carlos
 ; APPLICANT: Soto, Ana M.
 ; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Medien & Carroll, LLP
 ; STREET: 220 Montgomery Street, Suite 2200
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/769,746
 ; FILING DATE: 19-DEC-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Carroll, Peter G.
 ; REGISTRATION NUMBER: 32,837
 ; REFERENCE/DOCKET NUMBER: MBRI-02584
 ; TELEPHONE: (415) 705-8410
 ; TELEFAX: (415) 397-8338
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 585 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-769-746-2

Query Match 100.0%; Score 3103; DB 3; Length 585;
 Best Local Similarity 100.0%; Pred No. 9,3e-287;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKOLGENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADESAE 60
 |||||
 DB 1 DAHKEVAHRFKOLGENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADESAE 60
 |||||

QY 61 NCOKSLHTLFGDKLCTVATLRETYGENADCCAKOBERNECEFLQHKDDNPNLRLVRPEV 120
 |||||
 DB 61 NCOKSLHTLFGDKLCTVATLRETYGENADCCAKOBERNECEFLQHKDDNPNLRLVRPEV 120
 |||||

QY 121 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPELLFFAKYKAAFTCCQAADKAAACLLP 180
 |||||
 DB 121 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPELLFFAKYKAAFTCCQAADKAAACLLP 180
 |||||

QY 181 KLDELDRDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 240
 |||||
 DB 181 KLDELDRDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 240
 |||||

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDISSKLKECCCKPLEKSHCIAEVENDMPA 300
 |||||
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDISSKLKECCCKPLEKSHCIAEVENDMPA 300
 |||||

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYEARHPDYSVVVLLRLAKTYETTLK 360
 |||||
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYEARHPDYSVVVLLRLAKTYETTLK 360
 |||||

QY 361 CAAADPHECYAKVFDEFKPLVEEPQNLKONCELPEOLGEYKFONALLVRYTKKVPQVST 420
 |||||
 DB 361 CAAADPHECYAKVFDEFKPLVEEPQNLKONCELPEOLGEYKFONALLVRYTKKVPQVST 420
 |||||

QY 421 PTLVEVSRNLGVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 |||||
 DB 421 PTLVEVSRNLGVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 |||||

QY 481 LVNRRPCPSALEVDETYVPKFNATFTFHADICTLSEKEROIKKQATALVELVKKHKPRAT 540
 |||||
 DB 481 LVNRRPCPSALEVDETYVPKFNATFTFHADICTLSEKEROIKKQATALVELVKKHKPRAT 540
 |||||

QY 541 KEQLKAVMDDDFAAFVEKCKCKADKCTCFABEGKLVAAASQAALGL 585
 |||||
 DB 541 KEQLKAVMDDDFAAFVEKCKCKADKCTCFABEGKLVAAASQAALGL 585
 |||||

RESULT 4
 US-08-797-689-2
 ; Sequence 2, Application US/08797689
 ; Patent No. 5876969
 ; GENERAL INFORMATION:
 ; APPLICANT: Fleer, Reinhard
 ; APPLICANT: Fournier, Alain
 ; APPLICANT: Guittion, Jean-Dominique
 ; APPLICANT: Jung, Gerard
 ; APPLICANT: Yeh, Petrice
 ; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
 ; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
 ; NUMBER OF SEQUENCES: 36
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.
 ; STREET: 500 Arcola Road, 3C43
 ; CITY: Collegeville
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19426
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: System 7.1
 ; SOFTWARE: Word 5.1 (Patentin)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/797,689
 ; FILING DATE: 31-JAN-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/256,927
 ; FILING DATE: 28-JUL-1994
 ; APPLICATION NUMBER: FR 92/01064
 ; FILING DATE: 31-JAN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/FR93/00085
 ; FILING DATE: 28-JAN-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Smith Ph.D., Julie K.
 ; REGISTRATION NUMBER: P-38,619
 ; REFERENCE/DOCKET NUMBER: ST92006-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (610) 454-3839
 ; TELEFAX: (610) 454-3808
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 610 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-797-689-2

Query Match 100.0%; Score 3103; DB 2; Length 610;
 Best Local Similarity 100.0%; Pred No. 9,8e-287;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREKDLGEBNFKALVLIAPAOYLQOCPEDHVKLVNEVTEFAKTCVADESSE 60
DB 25 DAHSEVAHREKDLGEBNFKALVLIAPAOYLQOCPEDHVKLVNEVTEFAKTCVADESSE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNBEETLKKYLVEIARRHPYFYAPELLFPFAKRYKAAFTCCCAADKAACLLP 180
DB 145 DVMCTAFHDNBEETLKKYLVEIARRHPYFYAPELLFPFAKRYKAAFTCCCAADKAACLLP 204
QY 181 KLDELROEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRFPAEFAEVSCLKVTDLT 240
DB 205 KLDELROEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRFPAEFAEVSCLKVTDLT 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPPEKLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPPEKLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRJLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRJLAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEKPPLVEEPQNLKONCELFEQLGKVFQNALLVRYTKVPQVST 420
DB 385 CAADPHCEYAKVDFEKPPLVEEPQNLKONCELFEQLGKVFQNALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKKHPKAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLGKVGSKCKKHPKAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540
DB 505 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKERQIKKOTALVELVKGHPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 609

RESULT 5

US-08-256-938-2
; Sequence 2, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.0 (patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,938
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/01065
; FILING DATE: 31-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne

REGISTRATION NUMBER: 32,534
REFERENCE/DOCKET NUMBER: ST92007-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3817
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 783 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-256-938-2

Query Match 100.0%; Score 3103; DB 1; Length 783;
Best Local Similarity 100.0%; Pred. No. 1.4e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHREKDLGEBNFKALVLIAPAOYLQOCPEDHVKLVNEVTEFAKTCVADESSE 60
DB 25 DAHSEVAHREKDLGEBNFKALVLIAPAOYLQOCPEDHVKLVNEVTEFAKTCVADESSE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNBEETLKKYLVEIARRHPYFYAPELLFPFAKRYKAAFTCCCAADKAACLLP 180
DB 145 DVMCTAFHDNBEETLKKYLVEIARRHPYFYAPELLFPFAKRYKAAFTCCCAADKAACLLP 204
QY 181 KLDELROEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRFPAEFAEVSCLKVTDLT 240
DB 205 KLDELROEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRFPAEFAEVSCLKVTDLT 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPPEKLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPPEKLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRJLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRJLAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEKPPLVEEPQNLKONCELFEQLGKVFQNALLVRYTKVPQVST 420
DB 385 CAADPHCEYAKVDFEKPPLVEEPQNLKONCELFEQLGKVFQNALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKKHPKAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLGKVGSKCKKHPKAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540
DB 505 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKERQIKKOTALVELVKGHPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 609

RESULT 6

US-08-256-938-4
; Sequence 4, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA

COUNTRY: USA
 ZIP: 19426
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Macintosh
 OPERATING SYSTEM: System 7.1
 SOFTWARE: Word 5.0 (PatentIn)
 CURRENT APPLICATION DATA: US/08/256,938
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 92/01065
 FILING DATE: 31-JAN-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Goodman, Robanne
 REGISTRATION NUMBER: 32,534
 REFERENCE/DOCKET NUMBER: ST92007-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 454-3817
 TELEFAX: (610) 454-3808
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 787 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-256-938-4

Query Match 100.0%; Score 3103; DB 1; Length 787;
 Best Local Similarity 100.0%; Pred. No. 1.4e-286;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY : DAHKEVAHRFDLGEENFKALVLIATAQYLOQCPFDHVKLVNEVTEFAKTCVADSAE 60
 DB : DAHKEVAHRFDLGEENFKALVLIATAQYLOQCPFDHVKLVNEVTEFAKTCVADSAE 262
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLRVREV 120
 DB 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLRVREV 322
 QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 DB 323 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 382
 QY 181 KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSQRFPAEFAEVSCLVTDLT 240
 DB 383 KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSQRFPAEFAEVSCLVTDLT 442
 QY 241 VHTCECHGDLLECADRADLAKYICENODSISSKLKECCXPLLEKSHCIAEVENDEMPA 300
 DB 443 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVENDEMPA 502
 QY 301 DLPSLAADVESKDVCKNYAEAKQVFLGNFLYEYARRHPDYISVLLLLAKTYETTLK 360
 DB 503 DLPSLAADVESKDVCKNYAEAKQVFLGNFLYEYARRHPDYISVLLLLAKTYETTLK 562
 QY 361 CAADPHCEYAKVDFEFLVPEEPNLKQNCSELFQELGEYKFNALLVRYTKKVPQVST 420
 DB 563 CAADPHCEYAKVDFEFLVPEEPNLKQNCSELFQELGEYKFNALLVRYTKKVPQVST 622
 QY 421 PTLVEYSRNLGKVGSKCKHPKAMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 623 PTLVEYSRNLGKVGSKCKHPKAMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 682
 QY 481 LVNRRPCFSALEVDYTPKFNABETFTTHADICTLSEKERQIKQTALVELVGHKPKAT 540
 DB 683 LVNRRPCFSALEVDYTPKFNABETFTTHADICTLSEKERQIKQTALVELVGHKPKAT 742
 QY 541 KEOLKAVMDDDFAAFVEKCKKADDKETCFABEGKULVAASCAALGL 585
 DB 743 KEOLKAVMDDDFAAFVEKCKKADDKETCFABEGKULVAASCAALGL 787

RESULT 7

US-08-797-689-16
 Sequence 16, Application US/08797689
 Patent No. 5876969
 GENERAL INFORMATION:
 APPLICANT: Fleer, Reinhard
 APPLICANT: Fournier, Alain
 APPLICANT: Guittion, Jean-Dominique
 APPLICANT: Yeh, Patrice
 APPLICANT: Yeh, Gerard
 TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
 PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
 TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Road, 3C43
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19426
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Macintosh
 OPERATING SYSTEM: System 7.1
 SOFTWARE: Word 5.1 (PatentIn)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US 08/256,927
 FILING DATE: 28-JUL-1994
 APPLICATION NUMBER: FR 92/01064
 FILING DATE: 31-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/256,927
 FILING DATE: 28-JUL-1994
 APPLICATION NUMBER: FR 92/01064
 FILING DATE: 31-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/FR93/00085
 FILING DATE: 28-JAN-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith Ph.D., Julie K.
 REGISTRATION NUMBER: P-38,619
 REFERENCE/DOCKET NUMBER: ST92006-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 454-3839
 TELEFAX: (610) 454-3808
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 787 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-797-689-16
 Query Match 100.0%; Score 3103; DB 2; Length 787;
 Best Local Similarity 100.0%; Pred. No. 1.4e-286;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKEVAHRFDLGEENFKALVLIATAQYLOQCPFDHVKLVNEVTEFAKTCVADSAE 60
 DB 203 DAHKEVAHRFDLGEENFKALVLIATAQYLOQCPFDHVKLVNEVTEFAKTCVADSAE 262
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLRVREV 120
 DB 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLRVREV 322
 QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 DB 323 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 382
 QY 181 KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSQRFPAEFAEVSCLVTDLT 240
 DB 383 KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSQRFPAEFAEVSCLVTDLT 442

QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECECEKPLLEKSHCIAEVENDEMPA 300
DB 443 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECECEKPLLEKSHCIAEVENDEMPA 502
QY 301 DLPSLAADRVESKDVCKNYAEAKDVLGMFLYVYARRHPDYVWLLRLAKTYETTTLEKC 360
DB 503 DLPSLAADRVESKDVCKNYAEAKDVLGMFLYVYARRHPDYVWLLRLAKTYETTTLEKC 562
QY 361 CAAADPHECYAKVDFEFKPLVBEPPQNLIKONCELFEQLGEYKFNALLVRYTKVQVST 420
DB 563 CAAADPHECYAKVDFEFKPLVBEPPQNLIKONCELFEQLGEYKFNALLVRYTKVQVST 622
QY 421 PTLVEVSRNLGVGSKCKKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480
DB 623 PTLVEVSRNLGVGSKCKKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 682
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
DB 683 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 742
QY 541 KEOLKAVMDDFAAVFEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 743 KEOLKAVMDDFAAVFEKCKKADDKETCFABEGKKLVAASQAALGL 787

RESULT 8
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Angen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match 99.98; Score 3099; DB 1; Length 609;
Best Local Similarity 99.8; Pred. No. 2.4e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDLGEENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHFKDLGEENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQEPNERNECFLOHKDDNPNLRLVPEV 120

DB 85 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQEPNERNECFLOHKDDNPNLRLVPEV 144
QY 121 DVNCTAFHDNNEETFLKKYLYEYIARRHPYFAPBELLFFAKRYKAAFTTECCQAADKAACLLP 180
DB 145 DVNCTAFHDNNEETFLKKYLYEYIARRHPYFAPBELLFFAKRYKAAFTTECCQAADKAACLLP 204
QY 181 KLDELDRDEGKASSAKORLKCASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLT 240
DB 205 KLDELDRDEGKASSAKORLKCASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLT 264
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECECEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECECEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADRVESKDVCKNYAEAKDVLGMFLYVYARRHPDYVWLLRLAKTYETTTLEKC 360
DB 325 DLPSLAADRVESKDVCKNYAEAKDVLGMFLYVYARRHPDYVWLLRLAKTYETTTLEKC 384
QY 361 CAAADPHECYAKVDFEFKPLVBEPPQNLIKONCELFEQLGEYKFNALLVRYTKVQVST 420
DB 385 CAAADPHECYAKVDFEFKPLVBEPPQNLIKONCELFEQLGEYKFNALLVRYTKVQVST 444
QY 421 PTLVEVSRNLGVGSKCKKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480
DB 445 PTLVEVSRNLGVGSKCKKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 564
QY 541 KEOLKAVMDDFAAVFEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEOLKAVMDDFAAVFEKCKKADDKETCFABEGKKLVAASQAALGL 609

RESULT 9
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan:
; APPLICANT: Barr, Kathryn A.
; APPLICANT: Brierley, Russell A.
; APPLICANT: Thill, Gregory P.
; APPLICANT: Tschoop, Juerg F.
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
; TITLE OF INVENTION: PICHIA PASTORIS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433,037
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 91082
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELETYPE: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:

```

; SEQUENCE CHARACTERISTICS:
;   LENGTH: 609 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
; US-08-433-037-4

```

```

Query Match          99.9%; Score 3099; DB 1; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.4e-286;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGENFKALVLIAPAOYLQOCPEFHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRFKDGLGENFKALVLIAPAOYLQOCPEFHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDSLSHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECEFLQHKDDNPNLPRLVRPEV 120
DB 85 NCDSLSHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECEFLQHKDDNPNLPRLVRPEV 144
QY 121 DVMTCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 145 DVMTCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 204
QY 181 KLDELDRDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 205 KLDELDRDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 264
QY 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLKECCEKPLLEKSHGICIAEVENDENPA 300
DB 265 VHTCCGHDLLCADDRADLAKYICENQDSISSKLKECCEKPLLEKSHGICIAEVENDENPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYETVPKFEFNAETFTPHADICTLSEKERQIKKOTALVELVKGKPKAT 540
DB 505 LVNRRPCFSALEVDYETVPKFEFNAETFTPHADICTLSEKERQIKKOTALVELVKGKPKAT 564
QY 541 KEQLKAVMDFFAFAFVEKCKCKADDDKTCFABEGKKLVAAASQAALGL 585
DB 565 KEQLKAVMDFFAFAFVEKCKCKADDDKTCFABEGKKLVAAASQAALGL 609

```

```

RESULT 10
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CBA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-08-897-956A-2

```

```

Query Match          99.9%; Score 3099; DB 4; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.4e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGENFKALVLIAPAOYLQOCPEFHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRFKDGLGENFKALVLIAPAOYLQOCPEFHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDSLSHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECEFLQHKDDNPNLPRLVRPEV 120
DB 85 NCDSLSHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECEFLQHKDDNPNLPRLVRPEV 144
QY 121 DVMTCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 145 DVMTCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 204
QY 181 KLDELDRDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 205 KLDELDRDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 264
QY 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLKECCEKPLLEKSHGICIAEVENDENPA 300
DB 265 VHTCCGHDLLCADDRADLAKYICENQDSISSKLKECCEKPLLEKSHGICIAEVENDENPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYETVPKFEFNAETFTPHADICTLSEKERQIKKOTALVELVKGKPKAT 540
DB 505 LVNRRPCFSALEVDYETVPKFEFNAETFTPHADICTLSEKERQIKKOTALVELVKGKPKAT 564
QY 541 KEQLKAVMDFFAFAFVEKCKCKADDDKTCFABEGKKLVAAASQAALGL 585
DB 565 KEQLKAVMDFFAFAFVEKCKCKADDDKTCFABEGKKLVAAASQAALGL 609

```

```

RESULT 11
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075
; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04075
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids

```

```
/ TYPE: amino acid
/ STRANDEDNESS: unknown
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
PCT-US95-04075-3

Query Match      99.9%; Score 3099; DB 5; Length 609;
Best Local Similarity 99.8%; Pred. No. 2,4e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGGEENFKALVLIIFAQYLOQCPFDHVKLVNEVTEFAKTCVADESAB 60
DB 25 DAHSEVAHRFKDLGGEENFKALVLIIFAQYLOQCPFDHVKLVNEVTEFAKTCVADESAB 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 144
QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 145 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
DB 205 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 264
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLNQLVHEKTPVSDRVTKCCTES 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLNQLVHEKTPVSDRVTKCCTES 384
QY 361 CAADPHCEYAKVDFEKPFLVEBPQNLIKQNCSELFQOLGEYKFCQNALLVRYTKKVPQVST 420
DB 385 CAADPHCEYAKVDFEKPFLVEBPQNLIKQNCSELFQOLGEYKFCQNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNKGKSGCKCKHPEAKRMPCAEDYLSVVLNQLVHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNKGKSGCKCKHPEAKRMPCAEDYLSVVLNQLVHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETYPKFNAAETTFTHADICTLSEKERQIKKQATALVELVKHKPKAT 540
DB 505 LVNRRPCFSALEVDVETYPKFNAAETTFTHADICTLSEKERQIKKQATALVELVKHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKXKLVAAASQAALG 585
DB 565 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKXKLVAAASQAALG 609

RESULT 12
US-08-897-956A-3
; Sequence 3, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 978
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion polypeptide

US-08-897-956A-3
Query Match      99.7%; Score 3095; DB 4; Length 978;
Best Local Similarity 99.8%; Pred. No. 1.1e-285;
Matches 583; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGGEENFKALVLIIFAQYLOQCPFDHVKLVNEVTEFAKTCVADESAB 60
DB 212 DAHSEVAHRFKDLGGEENFKALVLIIFAQYLOQCPFDHVKLVNEVTEFAKTCVADESAB 272
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 272 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 331
QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 332 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 392
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
DB 392 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 451
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 452 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 511
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLNQLVHEKTPVSDRVTKCCTES 360
DB 512 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLNQLVHEKTPVSDRVTKCCTES 571
QY 361 CAADPHCEYAKVDFEKPFLVEBPQNLIKQNCSELFQOLGEYKFCQNALLVRYTKKVPQVST 420
DB 572 CAADPHCEYAKVDFEKPFLVEBPQNLIKQNCSELFQOLGEYKFCQNALLVRYTKKVPQVST 631
QY 421 PTLVEVSRLNKGKSGCKCKHPEAKRMPCAEDYLSVVLNQLVHEKTPVSDRVTKCCTES 480
DB 632 PTLVEVSRLNKGKSGCKCKHPEAKRMPCAEDYLSVVLNQLVHEKTPVSDRVTKCCTES 691
QY 481 LVNRRPCFSALEVDVETYPKFNAAETTFTHADICTLSEKERQIKKQATALVELVKHKPKAT 540
DB 692 LVNRRPCFSALEVDVETYPKFNAAETTFTHADICTLSEKERQIKKQATALVELVKHKPKAT 751
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKXKLVAAASQAALG 584
DB 752 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKXKLVAAASQAALG 795

RESULT 13
US-08-448-196A-3
; Sequence 3, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
```


ATTORNEY/AGENT INFORMATION:
 NAME: BROAD JR., ROBERT L.
 REGISTRATION NUMBER: 18,757
 REFERENCE/DOCKET NUMBER: XX/WFS-28402-2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 205-544-0021
 TELEFAX: 205-544-0258
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 585 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 US-08-448-196A-3

Query Match 99.7%; Score 3093; DB 1; Length 585;
 Best Local Similarity 99.7%; Pred. No. 8.3e-286;
 Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKOLGGEENFKALVLIAPAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHKSEVAHRFKOLGGEENFKALVLIAPAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNPRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNPRLVRPEV 120
 QY 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAAADKAACLIP 180
 DB 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAAADKAACLIP 180
 QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPKAEFAEVSKLVTDLT 240
 DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPKAEFAEVSKLVTDLT 240
 QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECKPLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECKPLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKQNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLEK 360
 DB 301 DLPSLAADFVESKDVCKQNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLEK 360
 QY 361 CAADHPHECYAKVDFEFKPLVEBPQNLIKQNCFLPQOLGEYKFQNALLVRYTKVPQVST 420
 DB 361 CAADHPHECYAKVDFEFKPLVEBPQNLIKQNCFLPQOLGEYKFQNALLVRYTKVPQVST 420
 QY 421 PTLVEVSRLNGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
 DB 421 PTLVEVSRLNGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
 QY 481 LVNRRPCFSALEVDVETVPKPFNAETFTPHADICTLSEKERQIKKOTALVELVKGHPKAT 540
 DB 481 LVNRRPCFSALEVDVETVPKPFNAETFTPHADICTLSEKERQIKKOTALVELVKGHPKAT 540
 QY 541 KEOLKAVMDDFAAVFEKCKKADDKETCFABEKGKLVAAASQAALGL 585
 DB 541 KEOLKAVMDDFAAVFEKCKKADDKETCFABEKGKLVAAASQAALGL 585

RESULT 14

US-08-984-176-1
 Sequence 1, Application US/08984176
 Patent No. 5948609
 GENERAL INFORMATION:
 APPLICANT: CARTER, DANIEL C
 APPLICANT: HO, JOSEPH X
 APPLICANT: RUKER, FLORIAN
 TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
 TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

FILE REFERENCE: 08/984.176
 CURRENT APPLICATION NUMBER: US/08/984.176
 CURRENT FILING DATE: 1997-12-03
 NUMBER OF SEQ ID NOS: 1
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 585
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-08-984-176-1

Query Match 99.7%; Score 3093; DB 2; Length 585;
 Best Local Similarity 99.7%; Pred. No. 8.3e-286;
 Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKOLGGEENFKALVLIAPAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHKSEVAHRFKOLGGEENFKALVLIAPAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNPRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNPRLVRPEV 120
 QY 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAAADKAACLIP 180
 DB 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAAADKAACLIP 180
 QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPKAEFAEVSKLVTDLT 240
 DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPKAEFAEVSKLVTDLT 240
 QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECKPLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECKPLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKQNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLEK 360
 DB 301 DLPSLAADFVESKDVCKQNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLEK 360
 QY 361 CAADHPHECYAKVDFEFKPLVEBPQNLIKQNCFLPQOLGEYKFQNALLVRYTKVPQVST 420
 DB 361 CAADHPHECYAKVDFEFKPLVEBPQNLIKQNCFLPQOLGEYKFQNALLVRYTKVPQVST 420
 QY 421 PTLVEVSRLNGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
 DB 421 PTLVEVSRLNGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
 QY 481 LVNRRPCFSALEVDVETVPKPFNAETFTPHADICTLSEKERQIKKOTALVELVKGHPKAT 540
 DB 481 LVNRRPCFSALEVDVETVPKPFNAETFTPHADICTLSEKERQIKKOTALVELVKGHPKAT 540
 QY 541 KEOLKAVMDDFAAVFEKCKKADDKETCFABEKGKLVAAASQAALGL 585
 DB 541 KEOLKAVMDDFAAVFEKCKKADDKETCFABEKGKLVAAASQAALGL 585

RESULT 15

US-08-448-196A-5
 Sequence 5, Application US/08448196A
 Patent No. 5780594
 GENERAL INFORMATION:
 APPLICANT: CARTER, DANIEL C.
 TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
 TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
 TITLE OF INVENTION: RELATED PROTEINS
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NASA
 STREET: MARSHALL SPACE FLIGHT CENTER
 CITY: HUNTSVILLE
 STATE: ALABAMA
 COUNTRY: USA
 ZIP: 35812

Job time : 31 secs

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/448,196A
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-448-196A-5

Query Match 79.2%; Score 2458.5; DB 1; Length 583;
Best Local Similarity 75.8%; Pred. No. 2,1e-225;
Matches 442; Conservative 70; Mismatches 70; Indels 1; Gaps 1;

QY 1 DAHKSEVAHRFKDLGERNFKALVLIAPAOYLQCPPEFHVHKLNVNEVTEFAKTCVADESAE 60
DB 1 DTHKSEIAHRFNLDGKHFKGLVLIAPSOYLQCPPEFHVHKLNVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKPEPERNECFLOHDDNENLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGELADCCCKPEPERNECFLOHDDNENLRLKPEP 119
QY 121 DVMTAFHNEETFLKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQADKAACLLP 180
DB 120 DAQCAAFQEDPDKFLGKYLVEVARRHPYFYCPPELLPHAEYKADFTCCPADDKLACLIP 179
QY 181 KLDELDRGKASSAKQKLCASQKQGERAPKAWAVARLSQRPKAEFAEVSKLVTDLCTK 240
DB 180 KLDALKERILLSSAKERLKCSSFQNGERAVKMSVARLSQKFPKADFAEVSKIVITDLCTK 239
QY 241 VHTCCCHDILLECADRADLAKYICENQDSISSKKECCCKPILKSHCIAEVNDEMPA 300
DB 240 VHKCCCHDILLECADRADLAKYICENQDSISSKKECCCKPILKSHCIAEVNDEMPA 299
QY 301 DLPSLAADFVESKDVCKNYABAKDVFGLMFLYEVARRHPYVSVVLLRLAKTVETILLEK 360
DB 300 DIPALAADFAEDKEIKCHYKDAKDVFLGTFLYEVARRHPYVSVVLLRLAKTVETILLEK 359
QY 361 CAADPHCEYKAVPDEPKPLVEEPQNLIKNCELPFQGLGEYKFNALLVRYTKVPOVST 420
DB 360 CAEADPPACRYTVDFQFTPLVEEPKSLVKKNCDLFEVGEYDFONALIVRYTKKAPQVSI 419
QY 421 PTLVEVSRNLGKVGCKCKPEAKRMCAEDYLSVLNQLCVLHEKTPVSDRVTKCCCTES 480
DB 420 PTLVEIGRTLGLKVGSRCKLPESERLPCSENHLALNRLCVLHEKTPVSEKITKCCCTDS 479
QY 481 LVNRRPCFSALEVEDYTVPKFEFNAETFTFHADICTLSEKERQIKKQATALVELVKHKPKAT 540
DB 480 LAERRPCFSALELEDGYPVKEFKETFTFHADICTLPEDEKQIKKQSAALVELVKHKPKAT 539
QY 541 KEQLKAVNDPFAAFVEKCKADDDKTCFAEEGKKLVAASQAAL 583
DB 540 KEQLKTVLGNFSFAFVAKCCGREDKEACFAEBGPKLVASSQLAL 582
```

Search completed: October 27, 2003, 15:36:27

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:35:09 ; Search time 72 Seconds
(without alignments)
1360.618 Million cell updates/sec

Title: US-09-832-929-18
Perfect score: 3103
Sequence: 1 DAHSEVAHRFKDLGEENFK.....TCFAEGGKLVAAQALGL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues
Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	585	10	US-09-929-552-2
2	3103	100.0	585	11	US-09-932-613-445
3	3103	100.0	585	11	US-09-984-010-26
4	3103	100.0	585	11	US-09-833-041-18
5	3103	100.0	585	12	US-10-153-604A-5
6	3103	100.0	585	12	US-09-833-117-18
7	3103	100.0	585	12	US-10-319-263-1
8	3103	100.0	585	12	US-10-319-263-2
9	3103	100.0	585	12	US-10-414-469-1
10	3103	100.0	585	12	US-10-414-469-2
11	3103	100.0	585	12	US-09-932-322-445
12	3103	100.0	585	12	US-10-413-831-1
13	3103	100.0	585	12	US-10-413-831-2
14	3103	100.0	585	14	US-10-153-064-5
15	3103	100.0	604	11	US-09-984-010-7

16	3103	100.0	609	11	US-09-919-039-370	Sequence 370, Appl
17	3103	100.0	609	12	US-10-153-604A-7	Sequence 7, Appl1
18	3103	100.0	609	12	US-10-365-623-23	Sequence 23, Appl
19	3103	100.0	609	14	US-10-153-064-7	Sequence 7, Appl1
20	3103	100.0	610	10	US-09-984-186-2	Sequence 2, Appl1
21	3103	100.0	610	15	US-10-237-667-2	Sequence 2, Appl1
22	3103	100.0	610	15	US-10-237-708-2	Sequence 2, Appl1
23	3103	100.0	610	15	US-10-237-866-2	Sequence 2, Appl1
24	3103	100.0	610	15	US-10-237-871-2	Sequence 2, Appl1
25	3103	100.0	610	15	US-10-237-624-2	Sequence 2, Appl1
26	3103	100.0	651	12	US-10-153-604A-133	Sequence 133, App
27	3103	100.0	651	14	US-10-153-064-133	Sequence 133, App
28	3103	100.0	652	12	US-10-153-604A-132	Sequence 132, App
29	3103	100.0	652	14	US-10-153-064-132	Sequence 132, App
30	3103	100.0	653	12	US-10-153-604A-131	Sequence 131, App
31	3103	100.0	653	14	US-10-153-064-131	Sequence 131, App
32	3103	100.0	656	12	US-10-153-604A-130	Sequence 130, App
33	3103	100.0	656	14	US-10-153-064-130	Sequence 130, App
34	3103	100.0	676	12	US-10-153-604A-127	Sequence 127, App
35	3103	100.0	676	12	US-10-153-604A-129	Sequence 129, App
36	3103	100.0	676	14	US-10-153-064-127	Sequence 127, App
37	3103	100.0	676	14	US-10-153-064-129	Sequence 129, App
38	3103	100.0	677	12	US-10-153-604A-125	Sequence 125, App
39	3103	100.0	677	14	US-10-153-064-125	Sequence 125, App
40	3103	100.0	680	12	US-10-153-604A-123	Sequence 123, App
41	3103	100.0	680	14	US-10-153-064-123	Sequence 123, App
42	3103	100.0	787	10	US-09-984-186-16	Sequence 16, Appl
43	3103	100.0	787	15	US-10-237-667-16	Sequence 16, Appl
44	3103	100.0	787	15	US-10-237-708-16	Sequence 16, Appl
45	3103	100.0	787	15	US-10-237-866-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids

```

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match      100.0%; Score 3103; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGLENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEPAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLGLENFKALVLIIFAQYLOQCPFEDHVKLVNEVTEPAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPRLVLRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPRLVLRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELFFAKRYKAAFTCCQAAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLT 240
Db 181 KDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLT 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVWLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVWLLRLAKTYETTTLEKC 360
QY 361 CAADPHCEYAKVDFEKFPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
Db 361 CAADPHCEYAKVDFEKFPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHTKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHTKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKPKAT 540
Db 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKPKAT 540
QY 541 KEOLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 3
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version: #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164

```

; FILING DATE: 19-DEC-1996
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 585 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
 US-09-984-010-26

Query Match 100.0%; Score 3103; DB 11; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2.4e-269;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKSEVAHRFKDGLGENFKALVLIAPAOYLQOCPPEDHVKLVNEVTEFAKTCVADESAR 60
 DB 1 DAHKSEVAHRFKDGLGENFKALVLIAPAOYLQOCPPEDHVKLVNEVTEFAKTCVADESAR 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEVSCLVTLTK 240
 DB 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEVSCLVTLTK 240
 QY 241 VHTECCHGDLLECCADRADLAKYICENQDSISSKKECCKEKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTECCHGDLLECCADRADLAKYICENQDSISSKKECCKEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVVLRLAKTYETTTLEKC 360
 DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVVLRLAKTYETTTLEKC 360
 QY 361 CAADPHCEYAKVDFBFLVEBPQNLIKQNCFLFQGLGEYFQNALVRYTKVPQVST 420
 DB 361 CAADPHCEYAKVDFBFLVEBPQNLIKQNCFLFQGLGEYFQNALVRYTKVPQVST 420
 QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFTTHADICTLSEKERQIKKQATALVELVGHKPKAT 540
 DB 481 LVNRRPCFSALEVDYTYVPKEFNAETFTTHADICTLSEKERQIKKQATALVELVGHKPKAT 540
 QY 541 KEQLKAWMDFFAFAFVKCKCKADDDKTCFAEBGKKLVAASQAALGL 585
 DB 541 KEQLKAWMDFFAFAFVKCKCKADDDKTCFAEBGKKLVAASQAALGL 585

RESULT 4

US-09-833-041-18
 ; Sequence 18, Application US/09833041
 ; Publication No. US20030125247A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen, Craig A.
 ; APPLICANT: Haseltine, William A.
 ; TITLE OF INVENTION: Albumin Fusion Proteins
 ; FILE REFERENCE: PF545
 ; CURRENT APPLICATION NUMBER: US/09/833,041
 ; PRIOR FILING DATE: 2001-04-12
 ; PRIOR APPLICATION NUMBER: 60/229,358
 ; PRIOR FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: 60/256,931
 ; PRIOR FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/199,384
 ; PRIOR FILING DATE: 2000-04-25
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 18
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-833-041-18

Query Match 100.0%; Score 3103; DB 11; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2.4e-269;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKSEVAHRFKDGLGENFKALVLIAPAOYLQOCPPEDHVKLVNEVTEFAKTCVADESAR 60
 DB 1 DAHKSEVAHRFKDGLGENFKALVLIAPAOYLQOCPPEDHVKLVNEVTEFAKTCVADESAR 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEVSCLVTLTK 240
 DB 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEVSCLVTLTK 240
 QY 241 VHTECCHGDLLECCADRADLAKYICENQDSISSKKECCKEKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTECCHGDLLECCADRADLAKYICENQDSISSKKECCKEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVVLRLAKTYETTTLEKC 360
 DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVVLRLAKTYETTTLEKC 360
 QY 361 CAADPHCEYAKVDFBFLVEBPQNLIKQNCFLFQGLGEYFQNALVRYTKVPQVST 420
 DB 361 CAADPHCEYAKVDFBFLVEBPQNLIKQNCFLFQGLGEYFQNALVRYTKVPQVST 420
 QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFTTHADICTLSEKERQIKKQATALVELVGHKPKAT 540
 DB 481 LVNRRPCFSALEVDYTYVPKEFNAETFTTHADICTLSEKERQIKKQATALVELVGHKPKAT 540
 QY 541 KEQLKAWMDFFAFAFVKCKCKADDDKTCFAEBGKKLVAASQAALGL 585
 DB 541 KEQLKAWMDFFAFAFVKCKCKADDDKTCFAEBGKKLVAASQAALGL 585

RESULT 5

US-10-153-604A-5
 ; Sequence 5, Application US/10153604A
 ; Publication No. US20030143191A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bell et al.
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
 ; FILE REFERENCE: PF556
 ; CURRENT APPLICATION NUMBER: US/10/153,604A
 ; CURRENT FILING DATE: 2002-05-24
 ; PRIOR APPLICATION NUMBER: 60/293,212
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens

US-10-153-604A-5

```
Query Match      100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREKDLGEENFKALVLIIFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHREKDLGEENFKALVLIIFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKRYKAFTBCCQAADKAACLLP 180
DB 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKRYKAFTBCCQAADKAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKCSLQKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
DB 181 KLDELDRDEGKASSAKQRLKCSLQKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVWLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVWLLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVDFEKPVEEPQNLIKONCELFEOQGEYKFNQALLVRYTKKVPQVST 420
DB 361 CAADAPHECYAKVDFEKPVEEPQNLIKONCELFEOQGEYKFNQALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGVSKCKCKHPEAKRMPCAEYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNGVSKCKCKHPEAKRMPCAEYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
```

RESULT 6

```
US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18
```

```
Query Match      100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREKDLGEENFKALVLIIFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHREKDLGEENFKALVLIIFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKRYKAFTBCCQAADKAACLLP 180
DB 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKRYKAFTBCCQAADKAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKCSLQKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
DB 181 KLDELDRDEGKASSAKQRLKCSLQKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVWLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVWLLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVDFEKPVEEPQNLIKONCELFEOQGEYKFNQALLVRYTKKVPQVST 420
DB 361 CAADAPHECYAKVDFEKPVEEPQNLIKONCELFEOQGEYKFNQALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGVSKCKCKHPEAKRMPCAEYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNGVSKCKCKHPEAKRMPCAEYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
```

RESULT 7

```
US-10-319-263-1
; Sequence 1, Application US/10319263
; Publication No. US2003018082CA1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISG007
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-319-263-1

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFDKDLGEENFKALVLIAFAQYLQOCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHKEVAHRFDKDLGEENFKALVLIAFAQYLQOCCPFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
DB 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLVEIARHPYFYAPELFFAKRYKAAFTTECCQAADKAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLVEIARHPYFYAPELFFAKRYKAAFTTECCQAADKAACLLP 180
QY 181 KLDELURDEGKASSAKORLKCASLOKFGERAFAKAVARLSORFPKAEFAEYSKLVTDLT 240
DB 181 KLDELURDEGKASSAKORLKCASLOKFGERAFAKAVARLSORFPKAEFAEYSKLVTDLT 240
QY 241 VHTTECHGDLLECADRADLAKYICENODS:SSKLKECCERPLLEKSHCIAEVENDEMPA 300
DB 241 VHTTECHGDLLECADRADLAKYICENODS:SSKLKECCERPLLEKSHCIAEVENDEMPA 300
QY 301 DPLSLAADPFVESKDVCKNYAEAKOVFLGMFLYVYARRHPDYSVLLRLAKTYETTLK 360
DB 301 DPLSLAADPFVESKDVCKNYAEAKOVFLGMFLYVYARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNL:KQNCELFEQLGEYKFNALLVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNL:KQNCELFEQLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCCCHPEAKMPCAEYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCCCHPEAKMPCAEYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKQIKKQATLALVELVHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKQIKKQATLALVELVHKPKAT 540
QY 541 KSQLKAVMDDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 541 KSQLKAVMDDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 8

US-10-319-263-2
; Sequence 2, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLTATION
US-10-319-263-2

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFDKDLGEENFKALVLIAFAQYLQOCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHKEVAHRFDKDLGEENFKALVLIAFAQYLQOCCPFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
DB 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLVEIARHPYFYAPELFFAKRYKAAFTTECCQAADKAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLVEIARHPYFYAPELFFAKRYKAAFTTECCQAADKAACLLP 180
QY 181 KLDELURDEGKASSAKORLKCASLOKFGERAFAKAVARLSORFPKAEFAEYSKLVTDLT 240
DB 181 KLDELURDEGKASSAKORLKCASLOKFGERAFAKAVARLSORFPKAEFAEYSKLVTDLT 240
QY 241 VHTTECHGDLLECADRADLAKYICENODS:SSKLKECCERPLLEKSHCIAEVENDEMPA 300
DB 241 VHTTECHGDLLECADRADLAKYICENODS:SSKLKECCERPLLEKSHCIAEVENDEMPA 300
QY 301 DPLSLAADPFVESKDVCKNYAEAKOVFLGMFLYVYARRHPDYSVLLRLAKTYETTLK 360
DB 301 DPLSLAADPFVESKDVCKNYAEAKOVFLGMFLYVYARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNL:KQNCELFEQLGEYKFNALLVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNL:KQNCELFEQLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCCCHPEAKMPCAEYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCCCHPEAKMPCAEYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKQIKKQATLALVELVHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKQIKKQATLALVELVHKPKAT 540
QY 541 KSQLKAVMDDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 541 KSQLKAVMDDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 9

US-10-414-469-1
; Sequence 1, Application US/10414469
; Publication No. US20030190691A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/414,469
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 09/806,247
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: PCT/US99/22905
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02

; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-414-469-1

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEBNFKALVLIAPQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEBNFKALVLIAPQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPAEFAEVSCLVTDLT 240
QY 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCECEKPLEKSHCIAEVENDEMPA 300
DB 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCECEKPLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVVLLRLAKTYETTTLEK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVVLLRLAKTYETTTLEK 360
QY 361 CAAADPHECYAKVDFEFPKLVPEEPQNLIKONCELFEQGEYKFNALLVRYTKKVPQVST 420
DB 361 CAAADPHECYAKVDFEFPKLVPEEPQNLIKONCELFEQGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKCLVAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKCLVAASQAALGL 585

RESULT 10

US-10-414-469-2
; Sequence 2, Application US/10414469
; Publication No. US2003019069A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; FILE REFERENCE: Kits
; CURRENT APPLICATION NUMBER: US/10/414,469
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 09/806,247
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: PCT/US99/22905

; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MOD RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLTATION
US-10-414-469-2

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEBNFKALVLIAPQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEBNFKALVLIAPQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPAEFAEVSCLVTDLT 240
QY 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCECEKPLEKSHCIAEVENDEMPA 300
DB 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCECEKPLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVVLLRLAKTYETTTLEK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVVLLRLAKTYETTTLEK 360
QY 361 CAAADPHECYAKVDFEFPKLVPEEPQNLIKONCELFEQGEYKFNALLVRYTKKVPQVST 420
DB 361 CAAADPHECYAKVDFEFPKLVPEEPQNLIKONCELFEQGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVHKHPRAT 540
DB 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVHKHPRAT 540
QY 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKCLVAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKCLVAASQAALGL 585

RESULT 11

US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.

APPLICANT: Potter, M. Daniel
APPLICANT: Fleming, Tony J.
APPLICANT: Ladner, Robert Charles
TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
FILE REFERENCE: Dlx-018.1 PCT: Dlx-018.1 US
CURRENT APPLICATION NUMBER: US/09/332,322
CURRENT FILING DATE: 2001-08-17
NUMBER OF SEQ ID NOS: 458
SOFTWARE: PatentIn version 3.1
SEQ ID NO 445
LENGTH: 585
TYPE: PRT
ORGANISM: HomoSapiens
US-09-932-322-445

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGKGFALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGKGFALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLPRVREV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLPRVREV 120
QY 121 DVMTAHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCCAADKAACLLP 180
DB 121 DVMTAHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCCAADKAACLLP 180
QY 181 KLDELDEGRASSAKORLKASLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLT 240
DB 181 KLDELDEGRASSAKORLKASLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLT 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILKSHCIAEVENDMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILKSHCIAEVENDMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLRLAKTYETTTLEKC 360
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKGKPKAT 540
DB 481 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKGKPKAT 540
QY 541 KEQLKAVMDFFAFAVEKCKADDKETCFABEGKLVAAASQAALGL 585
DB 541 KEQLKAVMDFFAFAVEKCKADDKETCFABEGKLVAAASQAALGL 585

RESULT 12
US-10-413-831-1
Sequence 1, Application US/10413831
Publication No. US20030194813A1
GENERAL INFORMATION:
APPLICANT: Lau Ph.D., Edward
APPLICANT: Winkler M.D., James V.
TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
TITLE OF INVENTION: Kits
FILE REFERENCE: ISC007
CURRENT APPLICATION NUMBER: US/10/413,831
CURRENT FILING DATE: 2003-04-15

PRIOR APPLICATION NUMBER: US/09/806,247
PRIOR FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 60/115,392
PRIOR FILING DATE: 1999-01-11
PRIOR APPLICATION NUMBER: 60/102,738
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: 09/165,926
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: 09/165,581
PRIOR FILING DATE: 1998-10-02
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 585
TYPE: PRT
ORGANISM: Homo sapiens
US-10-413-831-1

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGKGFALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGKGFALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLPRVREV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLPRVREV 120
QY 121 DVMTAHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCCAADKAACLLP 180
DB 121 DVMTAHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCCAADKAACLLP 180
QY 181 KLDELDEGRASSAKORLKASLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLT 240
DB 181 KLDELDEGRASSAKORLKASLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLT 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILKSHCIAEVENDMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILKSHCIAEVENDMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLRLAKTYETTTLEKC 360
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKGKPKAT 540
DB 481 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKGKPKAT 540
QY 541 KEQLKAVMDFFAFAVEKCKADDKETCFABEGKLVAAASQAALGL 585
DB 541 KEQLKAVMDFFAFAVEKCKADDKETCFABEGKLVAAASQAALGL 585

RESULT 13
US-10-413-831-2
Sequence 2, Application US/10413831
Publication No. US20030194813A1
GENERAL INFORMATION:
APPLICANT: Bar-Or M.D., David
APPLICANT: Lau Ph.D., Edward
TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
TITLE OF INVENTION: Kits

; FILE REFERENCE: ISC007
 ; CURRENT APPLICATION NUMBER: US/10/413,831
 ; CURRENT FILING DATE: 2003-04-15
 ; PRIOR APPLICATION NUMBER: US/09/856,247
 ; PRIOR FILING DATE: 2001-07-16
 ; PRIOR APPLICATION NUMBER: 60/115,392
 ; PRIOR FILING DATE: 1999-01-11
 ; PRIOR APPLICATION NUMBER: 60/102,738
 ; PRIOR FILING DATE: 1998-10-02
 ; PRIOR APPLICATION NUMBER: 09/165,926
 ; PRIOR FILING DATE: 1998-10-02
 ; PRIOR APPLICATION NUMBER: 09/165,581
 ; PRIOR FILING DATE: 1998-10-02
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: MOD_RES
 ; LOCATION: (1)..(585)
 ; OTHER INFORMATION: ACETYLATION
 US-10-413-831-2

Query Match 100.0%; Score 3103; DB 12; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2,4e-269;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 Db 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPRLVRPEV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPRLVRPEV 120
 Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 Qy 181 KLDELDEGKASSAKORLKCSLQKGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTG 240
 Db 181 KLDELDEGKASSAKORLKCSLQKGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTG 240
 Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILKSKSHCIAEVENDEMPA 300
 Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILKSKSHCIAEVENDEMPA 300
 Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
 Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
 Qy 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
 Db 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
 Qy 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
 Db 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
 Qy 481 LVNRRPCFSALEVDVETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 540
 Db 481 LVNRRPCFSALEVDVETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 540
 Qy 541 KEQLKAVMDDFAAFEVKCKCKADDKETCFABEGKKLVAASQAALGL 585
 Db 541 KEQLKAVMDDFAAFEVKCKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 14
 US-10-153-064-5
 ; Sequence 5, Application US/10153064

; Publication No. US20020142814A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bell et al.
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
 ; FILE REFERENCE: PF556
 ; CURRENT APPLICATION NUMBER: US/10/153,064
 ; CURRENT FILING DATE: 2002-05-24
 ; PRIOR APPLICATION NUMBER: 60/293,212
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-10-153-064-5
 Query Match 100.0%; Score 3103; DB 14; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2,4e-269;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 Db 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPRLVRPEV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPRLVRPEV 120
 Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 Qy 181 KLDELDEGKASSAKORLKCSLQKGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTG 240
 Db 181 KLDELDEGKASSAKORLKCSLQKGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTG 240
 Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILKSKSHCIAEVENDEMPA 300
 Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILKSKSHCIAEVENDEMPA 300
 Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
 Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
 Qy 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
 Db 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
 Qy 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
 Db 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
 Qy 481 LVNRRPCFSALEVDVETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 540
 Db 481 LVNRRPCFSALEVDVETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 540
 Qy 541 KEQLKAVMDDFAAFEVKCKCKADDKETCFABEGKKLVAASQAALGL 585
 Db 541 KEQLKAVMDDFAAFEVKCKCKADDKETCFABEGKKLVAASQAALGL 585
 RESULT 15
 US-09-984-010-7
 ; Sequence 7, Application US/09984010
 ; Publication No. US20030104578A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballance, David James
 ; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
 ; AND SERUM ALBUMIN
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP

STREET: 1300 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA: US/09/984,010
FILING DATE: 21-May-2002
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 09/091,873
FILING DATE: 25-JUN-1998
APPLICATION NUMBER: PCT/GB96/03164
FILING DATE: 19-DEC-1996
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 604 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-984-010-7

Search completed: October 27, 2003, 15:45:00
Job time : 73 secs

Query Match 100.0%; Score 3103; DB 11; Length 604;
Best Local Similarity 100.0%; Pred No. 2,5e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAHKSEVAHRFKDGLGENTFKALVLIAPQYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE	60
Db	20	DAHKSEVAHRFKDGLGENTFKALVLIAPQYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE	79
QY	61	NCCKSLHTLFGDKLCTVAT-RETYGEMADCCAKQPERNECFLOKHQDDNPMLRLVLRPEV	120
Db	80	NCCKSLHTLFGDKLCTVATRETYGEMADCCAKQPERNECFLOKHQDDNPMLRLVLRPEV	139
QY	121	DMVCTAFHNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP	180
Db	140	DMVCTAFHNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP	199
QY	181	KLDELDEGKASSAKORLKCASLOKEGERAFKAWAVARLSORFPAEAEVSKLVTDLT	240
Db	200	KLDELDEGKASSAKORLKCASLOKEGERAFKAWAVARLSORFPAEAEVSKLVTDLT	259
QY	241	VHTECHGDLLECADRADLAKYICENQDSISSKLECCCKP2LLEKSHCIAEVENDEMPA	300
Db	260	VHTECHGDLLECADRADLAKYICENQDSISSKLECCCKP2LLEKSHCIAEVENDEMPA	319
QY	301	DLPSLAADFVSKVCKVNAEAKDVLGMLFLEYARRHPDYVSVLLLLAKTYETITLKC	360
Db	320	DLPSLAADFVSKVCKVNAEAKDVLGMLFLEYARRHPDYVSVLLLLAKTYETITLKC	379
QY	361	CAAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFELQGEYKFNALLVRYTKKVPQVST	420
Db	380	CAAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFELQGEYKFNALLVRYTKKVPQVST	439
QY	421	PTLVESVRNLGKVGSKCKHPEAKRMPCAEDYLSVVNLQNLVHLKTPVSDRVTKCCTES	480
Db	440	PTLVESVRNLGKVGSKCKHPEAKRMPCAEDYLSVVNLQNLVHLKTPVSDRVTKCCTES	499
QY	481	LVNRRPCFSALEVDSTVVPKFNAAETFTSHADICTLSEKERQIKKOTALVELVHKHKPAT	540
Db	500	LVNRRPCFSALEVDSTVVPKFNAAETFTSHADICTLSEKERQIKKOTALVELVHKHKPAT	559
QY	541	KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAFEGKGLVAASQAALGL	585
Db	560	KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAFEGKGLVAASQAALGL	604

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:29:49 ; Search time 43 seconds
(without alignments)
1308.341 Million cell updates/sec

Title: US-09-832-929-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHFRKDLGEENFK.....TCFAEEGKKLVAASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:**

1: pirl:**

2: pirl:**

3: pirl:**

4: pirl:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	609	1	ABHUS
2	2942	94.8	600	2	A47331
3	2620	84.4	608	2	S57632
4	2475.5	79.8	607	1	ABHOS
5	2446.5	78.8	607	1	ABBSOS
6	2432.5	78.4	607	1	ABSHS
7	2426	78.2	608	1	ABRTS
8	2411.5	77.7	605	1	ABPGS
9	2387	76.9	609	2	JCS838
10	1861	60.0	453	2	A05139
11	1557.5	50.2	615	1	ABCHS
12	1253.5	40.4	609	2	JC4258
13	1249.5	40.3	609	1	PFHU
14	1242.5	40.0	609	1	FFGO
15	1205	38.8	607	1	ABXL72
16	1181.5	38.1	265	2	I46986
17	1175.5	37.9	608	1	ABXL68
18	1084	34.9	605	1	FWMS
19	1067	34.4	611	1	FRFT
20	1055	34.0	599	1	A44906
21	928.5	29.9	614	2	S59517
22	928	29.9	608	2	A53195
23	747.5	24.1	608	1	ABONS1
24	742.5	23.9	608	1	ABONS2
25	699	22.5	382	2	A37253
26	440.5	14.2	1423	1	S27941
27	386	12.4	474	1	VYHVD
28	385	12.4	476	1	VVRVD
29	372	12.0	472	1	A35327

RESULT: 1

ABHUS

serum albumin precursor [validated] - human

N/Alternate names: preproalbumin

N/Contains: Kinetensin

C/Species: Homo sapiens (man)

C/Date: 29-Jul-1981 #sequence revision 31-Jan-1997 #text change 17-Mar-2000

C/Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422;

R/Lawr, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houch, C.M.; Najarian, R.C.; See

Nucleic Acids Res. 9, 6103-6114, 1981

A/Title: The sequence of human serum albumin cDNA and its expression in Escherichia C

A/Reference number: A93743; MUID:82081882; PMID:6171778

A/Accession: A93743

A/Molecule type: mRNA

A/Residues: 1-419, 'K', 421-609 <LAW>

A/Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:g28591; PIDN:CA

R/Dugaiczky, A.; Law, S.W.; Dennison, O.E.

Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982

A/Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A/Reference number: A93936; MUID:82105994; PMID:6275391

A/Accession: A93936

A/Molecule type: mRNA

A/Residues: 1-120, 'G', 122-609 <DUG>

A/Cross-references: EMBL:V00494; NID:g28589; PIDN:CAA23753.1; PID:g28590

R/Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.

J. Biol. Chem. 261, 3244-3251, 1986

A/Title: The human albumin gene. Characterization of the 5' and 3' flanking regions a

A/Reference number: I39427; MUID:86140099; PMID:2419329

A/Accession: I39427

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-26 <URA>

A/Cross-references: GB:M13075; NID:g178330; PIDN:AAAS1688.1; PID:g553173

R/Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994

A/Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian fami

A/Reference number: I59286; MUID:94181575; PMID:8134387

A/Accession: I59286

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 282-290, 'KSRFDLQ' <WAT>

A/Cross-references: GB:S69192; NID:g546032; PIDN:AAB30282.1; PID:g546033

A/Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia

R/Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam

Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994

A/Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxy

A/Reference number: I59313; MUID:94294404; PMID:8022807

A/Accession: I59313

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 589-590, 'ALPFRVKNLLQVKLP' <MAD>

A/Cross-references: GB:S70799; NID:g547231; PIDN:AAB31177.1; PID:g547232

A:Note: this frame-shift variant is designated albumin Bazzano; four additional variants
 R:Menaya, J.; Parrilla, R.; Ayuso, M.S.
 submitted to the EMBL Data Library, March 1995
 A:Reference number: G08292
 A:Accession: G01747
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-120, 'G', 122-455 <MEN>
 A:Cross-references: ENBL:J22961; NID:9763428; P:DN:AAA64922.1; PID:g763431
 R:Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
 Biochem. J. 308, 321-325, 1995
 A:Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast *Kex2*
 A:Reference number: S55314; MUID:9275251; PMID:7755591
 A:Accession: S55314
 A:Molecule type: protein
 R:Meloun, B.; Moravek, L.; Kostka, V.
 FEBS Lett. 58, 134-137, 1975
 A:Title: Complete amino acid sequence of human serum albumin.
 A:Reference number: A91420; MUID:76187907; PMID:1225573
 A:Accession: A91420
 A:Molecule type: protein
 A:Residues: 25-117, 'EQ', 120-154, 'Q', 156-293, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-
 R:Roehr, U.; Spittler, G.; Tripiet, D.
 Justus Liebig's Ann. Chem. 9, 881-884, 1988
 A:Title: Isolation and structure elucidation of middle-molecular weight peptides from ur
 A:Reference number: S06422
 A:Note: this paper is in German, with an English abstract
 A:Accession: S06422
 A:Molecule type: protein
 A:Residues: 25-48 <RO>
 R:Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
 Arch. Biochem. Biophys. 305, 595-599, 1993
 A:Title: Mass spectrometric identification of modifications to human serum albumin treat
 A:Reference number: S36882; MUID:93384321; PMID:8373198
 A:Accession: S36882
 A:Molecule type: protein
 A:Residues: 45-67; 141-160; 311-337; 469-490; 570-581 <FIN>
 R:Kausler, E.; Spittler, G.
 Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
 A:Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol
 A:Reference number: S17599; MUID:92126241; PMID:1772598
 A:Accession: S17599
 A:Molecule type: protein
 A:Residues: 25-54; 354-357; 431-447 <KAU>
 A:Note: 49-Leu was also found
 R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mittra, S.P.
 J. Immunol. 143, 1680-1684, 1989
 A:Title: Structures of histamine-releasing peptides formed by the action of acid proteas
 A:Reference number: A45800; MUID:89341406; PMID:2474609
 A:Accession: A45800
 A:Molecule type: protein
 A:Residues: 166-173 <CAR>
 R:Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
 Biochem. Biophys. Res. Commun. 136, 983-988, 1986
 A:Title: The amino acid sequence of kretensin, a novel peptide isolated from pepts:n-tre
 A:Reference number: A03239; MUID:86242180; PMID:3087352
 A:Accession: A03239
 A:Molecule type: protein
 A:Residues: 166-173, 'L' <MO>
 R:Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, C.; Watkins, S
 Proc. Natl. Acad. Sci. U.S.A. 87, 8724-8725, 1990
 A:Title: Mutations in genetic variants of human serum albumin found in Italy.
 A:Reference number: A38255; MUID:91062352; PMID:2247440
 A:Accession: C38255
 A:Molecule type: protein
 A:Residues: 76-111 <GAL1>
 A:Accession: B38255
 A:Molecule type: protein
 A:Residues: 82-105, 'K', 107-110 <GAL2>
 A:Note: this variant is designated albumin Vibo Valentia
 A:Accession: A38255
 A:Molecule type: protein

A:Residues: 76-83, 'K', 85-106 <GAL3>
 A:Note: this variant is designated albumin Torino
 R:Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
 Eur. J. Biochem. 214, 437-444, 1993
 A:Title: The structural characterization and bilirubin-binding properties of albumin
 A:Reference number: S33298; MUID:93292504; PMID:8513793
 A:Accession: S33298
 A:Molecule type: protein
 A:Residues: 255-263, 'E', 265-281 <MIN1>
 A:Note: this variant is designated albumin Herborn
 R:Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Por
 Biochim. Biophys. Acta 1119, 232-238, 1992
 A:Title: Two albumins with identical electrophoretic mobility are produced by dif
 A:Reference number: S21078; MUID:92190239; PMID:1347703
 A:Accession: S21078
 A:Molecule type: protein
 A:Residues: 354-356, 'K', 358-378 <MIN2>
 A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is report
 R:He, X.M.; Carter, D.C.
 Nature 358, 209-215, 1992
 A:Title: Atomic structure and chemistry of human serum albumin.
 A:Reference number: A46756; MUID:92334427; PMID:1630489
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms
 R:Brown, J.R.; Shockley, P.; Behrens, P.Q.
 in 'The Chemistry and Physiology of the Human Plasma Proteins, Birg, D.H., ed., pp.23-
 A:Reference number: A94442
 A:Contents: annotation; three-dimensional structure and disulfide bonds
 R:Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.
 Collect. Czech. Chem. Commun. 42, 564-579, 1977
 A:Title: Disulfide bonds in human serum albumin.
 A:Reference number: A90930
 A:Contents: annotation; disulfide bonds
 R:Jacobsen, C.
 Biochem. J. 171, 453-459, 1978
 A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity bindi
 A:Reference number: A90299; MUID:78186630; PMID:656055
 A:Contents: annotation; bilirubin-binding site
 R:Petters, T.; Reed, R.G.
 in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, L., eds., 11-
 A:Title: Serum albumin: conformation and active sites.
 A:Contents: annotation; binding sites
 R:Harper, M.E.; Dugaiczky, A.
 Am. J. Hum. Genet. 35, 565-572, 1983
 A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein g
 A:Reference number: A90028; MUID:83279982; PMID:6192711
 A:Contents: annotation; gene position
 R:Walker, J.E.
 FEBS Lett. 66, 173-175, 1976
 A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic a
 A:Reference number: A46755; MUID:76257808; PMID:955075
 A:Contents: annotation
 A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic
 R:Bohney, J.P.; Fonda, M.L.; Felthoff, R.C.
 FEBS Lett. 298, 266-268, 1992
 A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-pho
 A:Reference number: A56294; MUID:92183881; PMID:1544460
 A:Contents: annotation
 A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described;
 atase activity
 C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthe
 i:ribin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (w
 C:Comment: A large number of variants of human serum albumin have been described.
 C:Genetics:
 A:Gene: GDB:ALB
 A:Cross-references: GDB:118990; OMIM:103600
 A:Map position: 4q11-4q13
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyr
 F.11-18/Dcmain: signal sequence #status predicted <SIG>
 F.19-24/Dcmain: propeptide #status experimental <PRO>
 F.25-609/Product: serum albumin #status experimental <MPT>
 F.29-202/Dcmain: serum albumin repeat homology <SAL>

F;166-174/Product: kinetensin #status experimental <KIP>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;217/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,413-592/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 3103; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 9.2e-198; Indels 0; Gaps 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRFKDLGEENFKALVLIAPFAQYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMTCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 145 DVMTCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAPKAWAVARLSQRFPAKAEFAEVSCLVTDLTK 240
DB 205 KLDELDEGKASSAKQRLKCSLQKFGERAPKAWAVARLSQRFPAKAEFAEVSCLVTDLTK 264
QY 241 VHTCCGGDLLECCADRADLAKYICENQDSISSKJKECCCKPILLEKSHCIAEVENDEMPA 320
DB 265 VHTCCGGDLLECCADRADLAKYICENQDSISSKJKECCCKPILLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYVSVLLRLAKTYETTLTK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYVSVLLRLAKTYETTLTK 384
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVHLKTPVSDRVTKCCTES 480
DB 445 PTLVEVRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVHLKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVVKHKPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVVKHKPKAT 564
QY 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609

RESULT 2
A47391
serum albumin precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding site
A:Reference number: A47391; MUID:93211971; PMID:8460152
A:Contents: B/B homozygote
A:Accession: A47391
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-600 <WAT>
A:Cross-references: GB:M90463; NID:g342294; PIDN:AAA36906.1; PID:g342295
A:Experimental source: liver
A>Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBI:P:128281)
C:Superfamily: serum albumin; serum albumin repeat homology
F;21-194/Domain: serum albumin repeat homology <SA1>
F;213-386/Domain: serum albumin repeat homology <SA2>

F;405-584/Domain: serum albumin repeat homology <SA3>

Query Match 94.8%; Score 2942; DB 2; Length 600;
Best Local Similarity 93.5%; Pred. No. 4e-187;
Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 17 DTHKSEVAHRFKDLGEENFKGLVLFVAFQYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 76
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 136
QY 121 DVMTCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 137 DVMTCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 196
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAPKAWAVARLSQRFPAKAEFAEVSCLVTDLTK 240
DB 197 KLDELDEGKASSAKQRLKCSLQKFGERAPKAWAVARLSQRFPAKAEFAEVSCLVTDLTK 256
QY 241 VHTCCGGDLLECCADRADLAKYICENQDSISSKJKECCCKPILLEKSHCIAEVENDEMPA 300
DB 257 VHTCCGGDLLECCADRADLAKYICENQDSISSKJKECCCKPILLEKSHCIAEVENDEMPA 316
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYVSVLLRLAKTYETTLTK 360
DB 317 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYVSVLLRLAKTYETTLTK 376
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
DB 377 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 436
QY 421 PTLVEVRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVHLKTPVSDRVTKCCTES 480
DB 437 PTLVEVRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVHLKTPVSDRVTKCCTES 496
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVVKHKPKAT 540
DB 497 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVVKHKPKAT 556
QY 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKKLVAASQAAL 583
DB 557 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKKLVAASQAAL 599

RESULT 3
S57632
serum albumin precursor - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
R:Grigori, C.; Hilger, C.; Grigori, F.; Hentges, F.
Gene 169, 295-296, 1996
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A:Reference number: JC4660; MUID:96194824; PMID:8647469
A:Accession: JC4660
A:Molecule type: mRNA
A:Residues: 1-608 <H12>
A:Cross-references: EMBL:X84842; NID:g886484; PIDN:CAA59279.1; PID:g886485
A:Experimental source: liver
C:Comment: This protein is the major protein component in plasma. It functions as a monomer.
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: liver; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRP>
F;25-608/Product: serum albumin #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>

Query Match 78.4%; Score 2432.5; DB 1; Length 607;


```
Best Local Similarity 75.0%; Pred. No. 2,1e-153; Mismatches 72; Indels 1; Gaps 1;
Matches 437; Conservative 73;

QY 1 DAHKEVAHRFKDGEENFKALVLIATFAQY:QQQPFEDHVKLVNVEVTERAKTCVADSAE 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25 DTHKEIAHRFRDLGEENFQGLVLIATFSQYLQQCPFDHVKLVNVEVTERAKTCVADSHA 84
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 85 GCDKSLHTLFGDELCKVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 143
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 121 DVMCTAFHDNEETFLKYLKYLRIARRHPYFYAPBL:FFAKRYKAAFTCCQADKAACLLP 180
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 144 DTLCEAFKADKKFKWGLYEVARRHPYFYAPBL:FFAKRYKAAFTCCQADKAACLLP 203
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 181 KLDELURDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSORFFKASFAEVSRLVTLT 240
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 204 KIDAMREKYLASSARQRLKASLOKFGGERAFKAWAVARLSORFFKASFAEVSRLVTLT 263
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 241 VHTCECHGDLLECADRADLAKYICENODS:ISSKLECCERPKLLEKSHCAEVENDEMPA 300
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 264 VHTCECHGDLLECADRADLAKYICENODS:ISSKLECCERPKLLEKSHCAEVENDEMPA 323
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 301 DLPSLAADPFVESKDYCKKYAAKDVFLGFMFLYEVARRHPDYVSVLLRLAKTYETLEKC 360
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 324 NLPLTADFAEDKEVKYQKAKVFLGFMFLYEVARRHPDYVSVLLRLAKTYETLEKC 383
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 361 CAADPHCEYAKVDFEPFLVEEPONL:IKONCELFEOQLGEYKFNALLVRYTKVPOVST 420
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 384 CAEDPHACIATVDFDKLHLVDEPNL:IKONCELFEOQLGEYKFNALLVRYTKVPOVST 443
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 421 PTLVSVSNL:GKVGSKCKHPKAKMPCAE:DLVSVLNOLCVLHEKTPVSDRVTKCCTES 480
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 444 PTLVSVSNL:GKVGSKCKHPKAKMPCAE:DLVSVLNOLCVLHEKTPVSDRVTKCCTES 503
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 481 LVNRRPCFSALEVDVTPKPFENATFTFHADICTLSEKQIKQKQALVELVKKHFKPAT 540
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 504 LVNRRPCFSDLTLDVTPKPFENATFTFHADICTLSEKQIKQKQALVELVKKHFKPAT 563
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 541 KEOLKAVMDFAAFVEKCKKADKDKTCFAEKGKLVASQAL 583
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 564 DEQLATVMEVFAVFDKCAADKDEGCFVLEGPKLVAHQAL 606
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
ABRTS
serum albumin precursor - rat
N:Alternate names: preproalbumin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-May-1979 #sequence revision 31-May-1979 #text change 22-Jun-1999
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R:Sargent, T. D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 76, 243-246, 1979
A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A:Reference number: A93872; MUID:81223722; PMID:7017712
A:Accession: A93872
A:Molecule type: mRNA
A:Residues: 1-608 <SAR>
A:Cross-references: GB:J00698; MID:g55627; PID:CAA24532.1; PID:g55628
R:Strauss, A. W.; Bennett, C. D.; Donohue, A. M.; Rodkey, J. A.; Alberts, A. W.
J. Biol. Chem. 252, 6846-6855, 1977
A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis
A:Reference number: A92211; MUID:77249657; PMID:893447
A:Note: Cleavages during protein maturation
A:Accession: A92211
A:Molecule type: protein
A:Residues: 1-38 <STR>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A:Reference number: A91946; MUID:78109429; PMID:564345
A:Accession: A91946
A:Molecule type: protein
```

```

262 VHKECHGDLLECADRADLAKYICENODTISTKLKECCDKPLLEKSHCIAEAKGDELPA 321
301 D2PSLAADPVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVWLLRLAKTYETTLKCC 360
322 DLNPLEHDFVEDKEVCCKNYEAKDVLGTFVLEYGSRHPDYSVSLLLRIAKIYEATLEDC 381
361 CAADPHCYAKVDFEDFKPLVVEEPNLKQNCLEFQELGEGYFQNALVRVTKKVPQVST 420
382 CAKEDPPACAYATVDFKQPLVDEPKNLKQNCLEFQELGEGYFQNALVRVTKKVPQVST 441
421 PTLVEVSRNLGVSKCKCHPBAKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
442 PTLVAVARKLGLVSRCKRPEERLSCAEDYLSVLNRLCVLHEKTPVSKVTKCCTES 501
481 LVNRRPCFSALVEDETYVPKPFNAETFFPHADICTLSEKERQIKQQTALVELVGHKPKAT 540
502 LVNRRPCFSALTPTDITYPKPFVEGTFPHADLCTLPDEKQIKQQTALVELVGHKPKAT 561
541 KEQLKAVNMDFAAFVEKCKCAADKKTCPAEGGKLV 576
562 BEQLATVLGNFAAFVQKCCAAPDHEACFAVEGPKFV 597

RESULT: 9
JC5838
albumin - Mongolian jird
C:Species: Meriones unguiculatus (Mongolian jird)
C>Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 19-May-2000
C:Accession: JC5838
R:Yoshida, K.; Seto-Onshima, A.; Sinohara, H.
DNA Res. 4, 351-354, 1997
A>Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in
A:Reference number: JC5838; MUID:98116663; PMID:9455485
A:Accession: JC5838
A:Molecule type: mRNA
A:Residues: 1-609 <YOS>
A:Cross-references: DBJ:AB006197; NID:92317277; PIDN:BAA2:765.1; P-D:g2317278
A:Experimental source: liver
C:Superfamily: serum albumin; serum albumin repeat homology
F:222-395/Domain: serum albumin repeat homology <SA2>

Query Match 76.9%; Score 2387; DB 2; Length 609;
Best Local Similarity 73.9%; Pred. No. 2.1e-150;
Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;

QY 2 AHKSEVARRKDLGLENFKALVLLAFQYLQCCPFEDHVKLVNVEYFAKTCVADESAEN 61
DB 27 AHKSEIARHYKDLGKYEYFGVLVLTFSQYLQCKSYEBHVKLVREYTFDFAKCAKDESAEN 86
QY 62 CDKSLHTLFGDKLCTVATRLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVRPEVD 121
DB 87 CDKSLHTLFGDKLCSLPNFGKYAEMADCCAKQEPERNECFLOHKDQNPPLPFKRAEPD 146
QY 122 VMCATFHNDETFLKYLVEYIARHPYFAPBELLFFAKRYKAAFTCCQAAKACALLPK 181
DB 147 AMCTAFQENAEAFMGHYLHEVARHPYFYGPELLYLDKTYAVLTTECCAADKGCALTPK 206
QY 182 LDELRECKASAKORLKCSLQKFGERAFAKAVARLSQRFPAEFAEYKSLVTDLTKV 241
DB 207 LDALKEKALVSAVRQRLKSSMKKFGERAFAKAVARMSQTFPNADFAETIKLATDITKV 266
QY 242 HTECCGDLLECADRADLAKYICENQDSISSKLKCECKPLLEKSHCIAEVENDEMPAD 301
DB 267 TQECCHGDLLECADRADLAKYICENQASISSKLOACCDKEMLOKSOCLAEVHDDMPAD 326
QY 302 LPSLAADPVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVWLLRLAKTYETTLKCC 361
DB 327 LPALTADVEDKDVCKNYAEAKDVLGTFLEYGSRHPDYSVSLLLRLAKYIYEATLEKCC 386
QY 362 AAADPHCYAKVDFEDFKPLVVEEPNLKQNCLEFQELGEGYFQNALVRVTKKVPQVSTP 421
DB 387 AEADPHACYGHVDFEDFKPLVVEEPNLKSNCELYKLGEGYFQNALVRVTKKAPQVSTP 446

```

```
QY 422 TLVEVRNLGKVGSCCKKHPEAKRMPCEADYLSVLNQLCVLHEKTPVSDRVTKCCCTESL 481
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 447 TLVEAARSLGRVTHCCALPEKRP-PCVEDYLSALNRLVCLLHEKTPVSEQVTKCCSGSL 506
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 482 VNRRPCFSALEVDYTVPKENASTFTFHADI:CTLSKERQ:KKQTALVELVKKHKVATK 54:
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 507 VERRPCFSALEVDYTVPKENASTFTFHANI:CTLPEKEKQEKQTAELVKKHQPATE 566
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 542 EOLKAVMDDFAAVFKCKKADKCTCPAEKGKLVAAASQAL 583
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 567 EQLKVMGDFAEFLSKCKQEDKEACFTTEGPKLVAESQKAL 608
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
RESULT 10
A05139
serum albumin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: A05139; I48638
R:Minghetti, P.P.; Law, S.W.; Dugaiczky, A.
Mol. Biol. Evol. 2, 347-358, 1985
A:Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudogenes
A:Reference number: A3055; MUID:88216123; PMID:2452956
A:Accession: A05139
A:Molecule type: mRNA
A:Residues: 1-418 <MIN>
A:Cross-references: GB:M16111; NID:G191764; PIDN:AAA37190.1; PID:G191765
R:Boccaccio, C.; Deschatrette, J.; Meunier-Rotival, M.
Gene 88, 181-186, 1990
A:Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the
A:Reference number: I48638; MUID:90269606; PMID:1971802
A:Accession: I48638
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 379-453 <BOC>
A:Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA31458.1; PID:G899334
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F:123-296/Domain: serum albumin repeat homology <SA2>
F:315-453/Domain: serum albumin repeat homology (fragment) <SA3>
Query Match 60.0%; Score 1861; DB 2; Length 453;
Best Local Similarity 72.2%; Pred. No. 9-5e-116;
Matches 327; Conservative 64; Mismatches 62; Indels 0; Gaps 0;
QY 75 CTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVLRPEVDVWCTAFHDEET 134
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 CAIPNLRENYGELADCT:KQEPERNECFLOHKDDNPNLPPPERPEABANCTSFKENPTTF 60
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 135 LKKYLYETARRHPYFYAPPELLFFAKRYKAAFTCCQAAKACALLPKLDELRLDGGKASSA 194
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 MGHYLYEHVARRHPYFYAPPELLYAYAEQYNEILTQCAEADKESCLTPKLDGKYEKALVSSV 120
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 195 KORLKACASLOKGERAFKAWAVARLSQRPKAEFAEVSCLKVTLTKVHTECHGDLLLECA 254
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 121 RORMKSSWQKGERAFKAWAVARLSQTPNADFAEITKLATD:TKYKNECHGDLLLECA 180
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 255 DDRADLAKYICENQDSISGKKECEKPLLEKSHCIAEVENDEMPADLPSSAAADFVSKD 3:4
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 181 DORAEALKVMENQATISSKLTQCCDKPKLLKKAHCLSEVEHDTMPADLPALAAADFVEDQE 240
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 315 VCKNYAEAKDVLGFLGYEYARRHPDYSVALLRLAKTYETTTLEKCCAAADPHCYAKVF 374
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 241 VCKNYAEAKDVLGFLGYEYARRHPDYSVLLRLAKTYETTTLEKCCAAADPHCYAKVF 300
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 375 DEFKPLVBPONLIRKONCFELQGEYKFNALLVRYTKKPVQVSTPTLVSEVSNLKGVG 434
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 301 AEFQPLVEBPKNLVKTNCDYELKGEYGFQNALVRYTKAPQVSTPTLVFAARNLGRVG 360
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 435 SKCKKHPEAKRMPCEADYLSVLNQLCVLHEKTPVSDRVTKCCCTESL:VNRRPCFSALEVD 494
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 361 TKCCTLPEDQRLPCVEDYLSALNRLVCLLHEKTPVSEHVTVCSSGSLVERPCFSALTVD 420
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

```
QY 495 ETVVPKEFNAETFTFHADI:CTLSKERQ:IKKQT 527
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 421 ETVVPKEFNAETFTFHSDICTLPEKQIKKQT 453
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
RESULT 11
ABCHS
serum albumin precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S15571; A05078; A13451
R:Cassady, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
A:Reference number: S15571
A:Accession: S15571
A:Molecule type: mRNA
A:Residues: 1-615 <CAS>
A:Cross-references: EMBL:X60688; NID:G63747; PIDN:CAA43098.1; PID:G63748
R:Hache, R.J.G.; Miskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A:Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein
A:Reference number: A05078; MUID:83161037; PMID:6187737
A:Accession: A05078
A:Molecule type: DNA
A:Residues: 1-28 <HAC>
R:Rosen, A.M.; Geller, D.M.
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977
A:Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.
A:Reference number: A13451; MUID:78019943; PMID:911327
A:Accession: A13451
A:Molecule type: protein
A:Residues: 19-23, 'M', 25-30 <ROS>
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds coppe
mones (weak bonds with these hormones promote their transfer across the membranes), t
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-26/Domain: propeptide #status predicted <PRO>
F:27-613/Product: serum albumin #status predicted <MAT>
F:32-206/Domain: serum albumin repeat homology <SA1>
F:225-398/Domain: serum albumin repeat homology <SA2>
F:417-596/Domain: serum albumin repeat homology <SA3>
F:30/Binding site: copper (His) #status predicted
F:80-89,102-118,117-128,152-197,196-205,228-274,273-281,293-307,306-317,344-389,388-3
```

```
Query Match 50.2%; Score 1557.5; DB 1; Length 615;
Best Local Similarity 46.7%; Pred. No. 1.6e-95;
Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;
QY 3 HKSEVAHRFDLGEENFKALVL11AFAYLQCCPFEDHVKLVNVEVTEFAKTCAVDESANEC 62
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 30 HKSEIAHRYNDLKESTFKAVAMITFAQLQRCVSEGLSKLVKDVVDLAQKAVANEDAPEC 89
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 63 DKSHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPR-LVRPEVD 121
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 90 SKPLPSIILDETCQVKLRDYSYGMADCCSKADPERNECFLSFKVQSPDFVQPYQRFASD 149
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 122 VMCTAPHONEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAFTCCQAAKACALLPK 181
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 150 VICQEQDNRVSVFLGHFYIYVARRHPFLYAPAILSFVDFEHALQSCCKESDVCACLDTK 209
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 182 LDELDEKQKASAKQRLKASLOKGERAFKAWAVARLSQRPKAEFAEVSCLKVTLTKV 241
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 210 EIVMREKAGVSKVQQYFCGILKQFGRVFOARQLIYLSQRYKAPFSEVSKFVHDSIGV 269
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 242 HTECHGDLLLECADRADLAKYICENQDSISGKKECEKPLLEKSHCIAEVENDEMPAD 301
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 270 HKECEGDMVECDMDMARMMNLCSSQDVFSGTKDKCEKPIVERSQCIMEAEFDEKPAD 329
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 302 LPSLAADFVESKDVCKNYAEAKDVLGFLGYEYARRHPDYSVLLRLAKTYETTTLEKCC 361
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

```

Db      330 LPSLVKEYIEDKEVCSFEAGHDAFMAEFVYVYSSRRHPBS:QLIMRIAKGVESLLEKCC 399
QY      362 AADPHCEYAKVDEPKPLVEEPONLIKONCELFEOQGEYKFKQNALLVYTKVPOVSTP 421
Db      390 KTDNPAECYANAQEQNLQHIKETQDVVVKNCNCDLHDHGEADFLKS:LIRYTKVPOVPTD 449
QY      422 TLVEVSRNLGKVGSKCKHPKAPCAEDYLSVLNQLCVLHETPVSVDRTVKCCTESL 481
Db      450 LLETGCKMTTIG:KCCQGLGEDRRMACSEGLYSIVHDCRQETTPINDNVSQCSQY 509
QY      482 VNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKATK 541
Db      510 ANRRPCFTAMGVDTKYVPPPEPDMFSFDEKLSAPAEEREVSQMKLLINLIKREQMT 569
QY      542 EQLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAAQAALGL 585
Db      570 EQIKTAGDTAMVDKCKKOSDINTCFGBEGANLIVOSRATLGI 613

RESULT 12
JC4258
alpha-fetoprotein precursor - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 20-Aug-1999
C:Accession: J04258
R:Rishio, H.; Gibbs, P.E.M.; Minghetti, P.P.; Zielinski, R.; Dugaiczky, A.
Gene 162, 213-220, 1995
A:Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to
A:Reference number: JC4258; MUID:96032345; PMID:7557431
A:Accession: J04258
A:Molecule type: DNA
A:Residues: 1-609 <NIS>
A:Cross-references: GB:U21916; NID:G84131; PIDN:AAA:641.1; P.D:G84:312
A:Comment: This protein is a plasma protein produced in the fetal and neonatal liver and
O similar properties and structure.
C:Genetics:
A:Gene: atp
A:Map position: 3p
A:Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551/4
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-609/Product: alpha-fetoprotein #status predicted <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:42,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.4%; Score 1253.5; DB 2; Length 609;
Best Local Similarity 40.1%; Pred. No. 2e-75;
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;

QY      3 HKSE-----VAHFPKDLGEENFALVLLIAPAVQLQCPEDHVKLVNVEYFAKTCVADE 57
Db      22 HRNEYGIASLDSYQCTAEINLTDLATIFPAQVQVQATYKEVSKMYKDALTAIERKPTGDE 81
QY      58 SAENCDKSLHTLFGDKLCTVATLRETYGEMADCAKOEERNECF:QHKDDNP-NLPLRV 116
Db      82 QSAGCLENQLPAFLBELCREKELEKYGH-SQCCSSQSEGRHNCFLAHKKPTASIPFPQ 140
QY      117 RPEVDVMTAFHNDNEFTLKKLYEIAHRRPYPAPELLFFAKRYKAAFTCCORADKAA 176
Db      141 VPEPVTSCBAYEEDRETFMKNFIYIARRHPFLYAPTILLWAARYDKIIPSCCKAENAVE 200
QY      177 CLLPKULDELDEKASAKQRLKASLQKRGFAFKAWAVARLSQRPFAEYVSKLVT 236
Db      201 CFQTKAATVTKELRESSLLNQHCACVMKNGRTFTQAITVTKLSQKFTKVFNETQKLV 260
QY      237 DLTQVTECCGDLLECCADRDALAKYICENODSISKKLKECKEPLEKSHCIAEVND 296
Db      261 DVAHVHEHCCRGVDLQDQGEKIMSYCSQQDTLSNKTTECKLTTLERGCITHAEND 320
QY      297 EMPADLPSLAADFVESKDVCKYAEAKDVLGMLFLEYARRHPDYSWLLLRANKTYET 356

```

```

Db      321 EKPEGLSPNLRFLGDDRFNOFSSGGEKNIFLASPVHEYSRRHPQLASVILURVAKGYOEL 380
QY      357 LEKCAAADPHCEYAKVDEPKPLVEEPONLIKONCELFEOQGEYKFKQNALLVYTKKVP 416
Db      381 LEKCFQTEPLECQDKGEELQKYIOESQALAKRSQGLFQKLGEYVLQNAFLVAYTKKAP 440
QY      417 QVSTPTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVLNQLCVLHETPVSVDRTVK 476
Db      441 QLTSSLMATIRKNAATAATCCQLSEDKLLACGEGAAADIIIGHLCIRHETTPVNPVGVC 500
QY      477 CTESLVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHK 536
Db      500 CTSSYANRRPCFSALVDVETVYPKAFSDDDKFIHKJUCQAGVALQTMKGEFLNLVKQK 560
QY      537 PKATKEQLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAAQAALGL 585
Db      561 PQTTEEQLAEVIADFSGLLEKCKCQGEVCFABEGQKLISKTRAAALGV 609

RESULT 13
FPHU
alpha-fetoprotein precursor [validated] - human
N:Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 08-Dec-2000
C:Accession: A26624; S37655; A93961; A91497; A23699; A61480; A90624; A90757; A93042; A
R:Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugaiczky, A.
Biochemistry 26, 1332-1343, 1987
A:Title: Structure, polymorphism, and novel repeated DNA elements revealed by a complete
A:Reference number: A26624; MUID:87185438; PMID:2436661
A:Accession: A26624
A:Molecule type: DNA
A:Residues: 1-609 <GIB>
A:Cross-references: GB:M16110; NID:G773678; PIDN:AAB58754.1; P.D:G178236
R:McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tilghman, S.; Krume
Hum. Mol. Genet. 2, 379-384, 1993
A:Title: A G->A substitution in an HNF I binding site in the human alpha-fetoprotein
A:Reference number: S37655; MUID:93278385; PMID:7684942
A:Accession: S37655
A:Molecule type: DNA
A:Residues: 1-28 <MCV>
A:Cross-references: EMBL:Z19532; NID:Q28527; PIDN:CAA79592.1; PID:G28528
A:Note: the authors translated the codon TAT for residue 26 as Thr
R:McIninga, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.
Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983
A:Title: Primary structures of human alpha-fetoprotein and its mRNA.
A:Reference number: A93961; MUID:83273664; PMID:6192439
A:Accession: A93961
A:Molecule type: mRNA
A:Residues: 1-609 <MCV>
A:Cross-references: GB:J00077; NID:G311348; PIDN:CAA24758.1; PID:G31351
R:Beattie, W.G.; Dugaiczky, A.
Gene 20, 415-422, 1982
A:Title: Structure and evolution of human alpha-fetoprotein deduced from partial sequen
A:Reference number: A91497; MUID:83158778; PMID:6187626
A:Accession: A91497
A:Molecule type: mRNA
A:Residues: 429-536 <BEA>
A:Cross-references: GB:J00076
R:Pucci, P.; Scilliano, R.; Maiorni, A.; Marino, G.; Tecce, M.F.; Ceccarini, C.; Terra
Biochemistry 30, 5061-5066, 1991
A:Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.
A:Reference number: A23699; MUID:91242409; PMID:1709810
A:Accession: A23699
A:Molecule type: protein
A:Residues: 19-45;60-97;102-107;122-184;187-249;255-489;507-609 <PUC>
R:Tecce, M.F.; Terrana, B.; Giuliani, M.M.; Ceccarini, C.
J. Nucl. Med. Allied Sci. 34, 213-216, 1990
A:Title: Characterization of in vitro expressed human alpha-fetoprotein as highly repr
A:Reference number: A61480; MUID:91225826; PMID:1709209
A:Accession: A61480
A:Molecule type: protein

```



```

Db 141 VPEPVTCEAYEEDRETFYKMKFIETARRHPFLYAPTILLWAARYDKIIPSCCKEENAVE 200
QY 177 CLLPKLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSORFPKAEAEVSKVT 236
Db 201 CFQTKAATVTKELRESSLLNQHCAMVKNFTGTFOAIVTKLSQKFTKVNTEIQKVL 260
QY 237 DLTQVHTCECHGDJLECCADRDADLAKYICENQDSISSKLECKECPLEKSHCIAEVDN 296
Db 261 DVAHVHEHCCRGVDLCLQDGEKIMSYICSQDSTLKNKITECCCKTLTGERGQCIHAEND 320
QY 297 EMPADLPSLAADPVESKVDKCNVAEAKDVLGMLFELYEYARRHPDYVSVLLBLAKTYET 356
Db 321 EKPEGLSPNLRFLGDRDFNQFSSGKRNIFLAFVHEYSRRHPQLAVSVILRVAKGYOEL 380
QY 357 LEKCAAAADPHECYAKVDFEFKPLVEEPQNLKONCELFQELGEYKFNALLVRYTKVP 416
Db 381 LEKCFOTENPLECODKGEELQKYIQESQALAKESCLGKQGLGEYQLQNAFLVATKXAP 440
QY 417 QVSTPTLVEVSRNLGKVGSKCKRHPKAPKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTK 476
Db 441 QLTSSSLMAITRMAATAATCCQLSEDKLACGEGAADIIGHLCIRHEMTVPNPGVQC 500
QY 477 CTESLVNRRPCFSALEVDYVVPKFNABTFTFHADICTLSEKEROIKKOTALVELVKHK 536
Db 501 CTSSYANRRPCFSSLVVDYVVPFASDDKFIHKDLQCAQGVALQTMKQEFNLNLVKOK 560
QY 537 PKATKEQLKAMVDPAFAFVEKCKCKADDKETCFAPKEGKLVAAQAALGL 585
Db 561 POITEQLETVIADFSGLEKCCQGOEQVCFAPKEGKLVSKTRTALGV 609

RESULT 15
ABXL72
74K albumin precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: B41682; S02693; A05288
R:Moskatis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
Mol. Endocrinol. 3, 464-473, 1989
A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic acid during development.
A:Reference number: A41682; MUID:89313786; PMID:2747653
A:Accession: B41682
A:Molecule type: mRNA
A:Residues: 3-607 <MOS>
A:Cross-references: GB:M2:442; NID:G213930; PID:AAA49637.1; PID:G213931;
R:Schorpp, M.; Doebebling, U.; Wagner, U.; Kyffel, G.J.
J. Mol. Biol. 199, 83-93, 1988
A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Deletion during development.
A:Reference number: S02692; MUID:88172470; PMID:2451026
A:Accession: S02693
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-48 <SCH>
A:Cross-references: EMBL:Z26826
R:Wolfe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J.
Eur. J. Biochem. 146, 489-496, 1985
A:Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabilization of mRNA.
A:Reference number: A05288; MUID:85126974; PMID:3971963
A:Accession: A05288
A:Molecule type: mRNA
A:Residues: 459-502, 'L', 504-557 <MOL>
A:Cross-references: GB:M28276
A>Note: the authors translated the codon TAT for residue 63 as Thr
C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper, mones (weak bonds with these hormones promote their transfer across the membranes), thyroglobulin, and retinol.
C:Genetics:
A:Introns: 27/1
A:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>

```

```

F:25-607/Product: 74K serum albumin #status predicted <MAT>
F:32-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:30/Binding site: copper (His) #status predicted
F:80-88/101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-3:
F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 38.8%; Score 1205; DB 1; Length 607;
Best Local Similarity 39.3%; Pred. No. 3.3e-72;
Matches 227; Conservative 108; Mismatches 239; Indels 4; Gaps 2;

QY 3 HKSEVAHRFKDLGEENPKALVLIAPAYIQOQPFEDHVKLVNEVTEFAKTCVADSAENC 62
Db 30 HHKHIADVYVYALTERTFKGLTIAIVSQNLQKSLSELSKLVNEINDFAKSCINDRTPC-C 88
QY 63 DKSLSHTLFGKLCVATLRETYGEMADCCAKOPBERNECFLOHKDDNPMLPRLVRPEVDV 122
Db 89 EKPVGTLFPDKLCADPAVGNYVYEWSEKCAKQDPERAQCFAHRDHEHT---SIKPEPEE 145
QY 123 MCTAFHDNEETFLKXYLYETARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAALLPKL 182
Db 146 TCKLLKEHPDOLLSAFTHEEARNHPDLYPPAVLALTQYHKLAECCEBEDKEKCFSEKM 205
QY 183 DELRDEGKASSAKORLKASLQKGERAFKAWAVARLSORFPKAEAEVSKLVTDLTQVH 242
Db 206 KQLMKQSHSIEDKQHFCWILDNFPKVLKALNARVSHRYPKAEFKLAHNTVEVTHFI 265
QY 243 TBCCHGDLLECCADRDADLAKYICENQDSISSKLECKECPLEKSHCIAEVDNENPADL 302
Db 266 KDCCHDDMFECMTERLELTHTCQHKDELSSKLEKCNIPLLERTYCVITLNDNDVPAEL 325
QY 303 PSLAADPVESKVDKCNVAEAKDVLGMLFELYEYARRHPDYVSVLLBLAKTYETL 362.
Db 326 SQPTETEDPHVCEKYAENNEVFLGRYUHAVSRKFEQELSEQFLQSAREYBSJLNKCK 385
QY 363 AADPHECYAKVDFEFKPLVEEPQNLKONCELFQELGEYKFNALLVRYTKVPQVSTPT 422
Db 386 TDNPPECYKGDADRFEAKERFAYLKQNCIDILHEHGEVLFENELLIRYTKMPQVSD 445
QY 423 LVEVSRNLGKVGSKCKRHPKAPKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTESLV 482
Db 446 LIGIAHQMADIGEHCCCAVPEQMPCAEGDLTILIGMCEKQKTTFINNHVAHCCTDSYS 505
QY 483 NRRPCFSALEVDYVVPKFNABTFTFHADICTLSEKEROIKKOTALVELVKHKPATKE 542
Db 506 GMRSCFTALGPDEYVPPVTDTHFDKICTANDKEKQHIKQKFLVKLIKVSFKLEKN 565
QY 543 QLKAVMDDFAAFVEKCKCKADDKETCFAPKEGKLVAAASQ 580
Db 566 HIDECSAEFLKMGVKCCTADEHQPCFDEKPVLIHCQ 603

```

Search completed: October 27, 2003, 15:35:50

Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:22:38 / Search time 25 Seconds
(without alignments)
2100.425 Million cell updates/sec

Title: US-09-832-929-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHRFKDLGEENFK.....TCFAEKGKLYAASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3103	100.0	609	1	ALBU_HUMAN
2	2942	94.8	600	1	ALBU_MACMU
3	2620	84.4	608	1	ALBU_FELCA
4	2562	82.6	608	1	ALBU_CANFA
5	2475.5	79.8	607	1	ALBU_HORSE
6	2450.5	79.0	607	1	ALBU_BOVIN
7	2446	78.8	608	1	ALBU_RABIT
8	2432.5	78.4	607	1	ALBU_SHEEP
9	2426	78.2	608	1	ALBU_RAT
10	2411.5	77.7	605	1	ALBU_P.G
11	2387	76.9	609	1	ALBU_MERUN
12	2378	76.6	608	1	ALBU_MOUSE
13	1557.5	50.2	615	1	ALBU_CHICK
14	1253.5	40.4	609	1	FETA_PANTR
15	1249.5	40.3	609	1	FETA_HUMAN
16	1242.5	40.0	609	1	FETA_GORGO
17	1205	38.8	607	1	ALB2_XENLA
18	1200	38.7	609	1	FETA_HORSE
19	1164.5	37.5	606	1	ALB1_XENLA
20	1084	34.9	605	1	FETA_MOUSE
21	1067	34.4	611	1	FETA_RAT
22	1055	34.0	599	1	AFAM_HUMAN
23	944	30.4	611	1	AFAM_MOUSE
24	928	29.9	608	1	AFAM_RAT
25	747.5	24.1	608	1	ALB1_SALSA
26	742.5	23.9	608	1	ALB2_SALSA
27	699	22.5	382	1	ALBU_RANCA
28	440.5	14.2	1423	1	ALBU_PETMA
29	386	12.4	474	1	VTDB_HUMAN
30	381	12.3	476	1	VTDB_RAT
31	378	12.2	476	1	VTDB_RABIT
32	372	12.0	472	1	VTDB_MOUSE
33	151.5	4.9	1605	1	RRB1_MOUSE

ALIGNMENTS

RESULT 1

ID	ALBU_HUMAN	STANDARD	PR	609 AA
AC	P02758; Q95574; Q13140; Q9P157; Q9P117; Q9JHS3; Q9UJZ0;			
Dr	21-JUL-1986 (Rel. 01, Created)			
Dr	01-APR-1990 (Rel. 14, Last sequence update)			
Dr	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Serum albumin precursor.			
GN	ALB.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
CX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86196112; PubMed=3009475;			
RA	Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,			
RA	Beatlie W.G., Dugaiczky A.;			
RT	"Molecular structure of the human albumin gene is revealed by			
RT	nucleotide sequence within q11-22 of chromosome 4.";			
RL	J. Biol. Chem. 261:6747-6757(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND VARIANT LYS-420.			
RX	MEDLINE=82081882; PubMed=6171778;			
RA	Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,			
RA	Najarian R.C., Seeburg P.H., Wior K.L.;			
RT	"The sequence of human serum albumin cDNA and its expression in E.			
RT	coli.";			
RN	Nucleic Acids Res. 9:6103-6114(1981).			
RP	[3]			
RX	SEQUENCE FROM N.A., AND VARIANT GLY-121.			
RA	MEDLINE=82105994; PubMed=6275391;			
RA	Dugaiczky A., Law S.W., Dennison O.E.;			
RT	"Nucleotide sequence and the encoded amino acids of human serum			
RT	albumin mRNA.";			
RN	Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).			
RP	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RA	Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;			
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).			
RC	TISSUE=Fetal liver;			
RA	Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,			
RA	Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;			
RT	"Functional prediction of the coding sequences of 121 new genes			
RT	deduced by analysis of cDNA clones from human fetal liver.";			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.			
RA	Huang M.C., Wu H.T.;			
RT	"The cDNA sequences of human serum albumin.";			
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.			
RN	[7]			
RP	SEQUENCE FROM N.A.			

34	144.5	4.7	8797	1	SNE1_HUMAN	Genf91 homo sapien
35	138.5	4.5	1410	1	RRB1_HUMAN	Q9P2e9 homo sapien
36	133.5	4.3	1391	1	MST2_DROHY	Q98696 drosophila
37	132.5	4.3	2230	1	GOG4_HUMAN	Q13439 homo sapien
38	129.5	4.2	1972	1	MYHB_HUMAN	P35749 homo sapien
39	129	4.2	3210	1	CENF_HUMAN	P49454 homo sapien
40	128	4.1	1005	1	RA50_METJA	Q58718 methanococc
41	126.5	4.1	1972	1	MYHB_RABIT	P35748 oryctolagus
42	126	4.1	1189	1	SMC2_CHICK	Q90988 gallus gall
43	126	4.1	3259	1	GIAN_HUMAN	Q14789 homo sapien
44	125	4.0	3038	1	TRIO_HUMAN	O75962 homo sapien
45	124.5	4.0	1790	1	USO1_YEAST	P25386 saccharomyc

RC TISSUE=Liver, and Skeletal muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8].
 RP SEQUENCE OF 25-609.
 RX MEDLINE=76187907; PubMed=1225573;
 RA Meloun B., Moravsek L., Kostka V.;
 RT "Complete amino acid sequence of human serum albumin.";
 RL FEBS Lett. 58:134-137(1975).
 RN [9].
 RP SEQUENCE OF 25-609.
 RX Brown J.R., Shockley P., Behrens P.Q.;
 RL (In) Bing D.H. (eds.);
 RL The chemistry and physiology of the human plasma proteins, pp.23-40,
 RL Pergamon Press, New York (1979).
 RN [10].
 RP SEQUENCE OF 1-455 FROM N.A.
 RC TISSUE=Liver;
 RA Menaya J., Parrilla R., Ayuso M.S.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [11].
 RP SEQUENCE OF 1-26 FROM N.A.
 RX MEDLINE=86140099; PubMed=2419329;
 RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
 RT "The human albumin gene. Characterization of the 5' and 3' flanking
 regions and the polymorphic gene transcripts.";
 RL J. Biol. Chem. 261:3244-3251(1986).
 RN [12].
 RP SEQUENCE OF 222-229.
 RX MEDLINE=76257808; PubMed=955075;
 RA Walker J.E.;
 RT "Lysine residue 199 of human serum albumin is modified by
 acetylsalicylic acid.";
 RL FEBS Lett. 66:173-175(1976).
 RN [13].
 RP SEQUENCE OF 25-44 AND 480-499.
 RC TISSUE=Heart;
 RX MEDLINE=95203287; PubMed=7895732;
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
 RT "The human myocardial two-dimensional gel protein database: update
 1994.";
 RL Electrophoresis 15:1459-1465(1994).
 RN [14].
 RP BISULFIDE BONDS.
 RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;
 RT "Disulfide bonds in human serum albumin.";
 RL Collect. Czech. Chem. Commun. 42:564-579(1977).
 RN [15].
 RP BILIRUBIN-BINDING SITE.
 RX MEDLINE=78186630; PubMed=656055;
 RA Jacobsen C.;
 RT "Lysine residue 240 of human serum albumin is involved in high-
 affinity binding of bilirubin.";
 RL Biochem. J. 171:453-459(1978).
 RN [16].
 RP VARIANT CANTERBURY ASN-337.
 RX MEDLINE=87157744; PubMed=3828358;
 RA Brennan S.O., Herbert P.;
 RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
 domain of serum albumin.";
 RL Biochim. Biophys. Acta 912:191-197(1987).
 RN [17].
 RP VARIANTS NAG-2 AND NAG-3.
 RX MEDLINE=88068523; PubMed=3479777;
 RA Takashashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
 RA Satoh C., Neel J.V.;
 RT "Amino acid substitutions in inherited albumin variants from
 Amerindian and Japanese populations.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
 RN [18].
 RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
 RX MEDLINE=89345611; PubMed=2762316;
 RA Arai K., Madison J., Shimizu A., Putnam F.W.;
 RA Neel J.V., Sakurabayashi I., Putnam F.W.;
 RT "Point substitutions in Japanese alloalbumins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6032-6036(1989).
 RN [19].
 RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
 RX MEDLINE=90115905; PubMed=2404284;
 RA Arai K., Madison J., Shimizu A., Putnam F.W.;
 RT "Point substitutions in albumin genetic variants from Asia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
 RN [20].
 RP DESCRIPTION OF VARIANT REDHILL.
 RX MEDLINE=90115852; PubMed=2104980;
 RA Brennan S.O., Wyles T., Peach R.J., Donaldson D., George P.M.;
 RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
 human serum albumin whose precursor has an aberrant signal peptidase
 cleavage site.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
 RN [21].
 RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
 RX MEDLINE=91062352; PubMed=2247440;
 RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
 RA Watkins S., Putnam F.W.;
 RT "Mutations in genetic variants of human serum albumin found in
 Italy.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
 RN [22].
 RP VARIANT VENEZIA.
 RX MEDLINE=91296740; PubMed=2068071;
 RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
 RA Minchiotti L., Putnam F.W.;
 RT "A donor splice mutation and a sirg-e-base deletion produce two
 carboxyl-terminal variants of human serum albumin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
 RN [23].
 RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
 RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
 RX MEDLINE=92052189; PubMed=1946412;
 RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
 RA Matsuda Y.-I., Amaki I., Putnam F.W.;
 RT "Genetic variants of serum albumin in Americans and Japanese.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
 RN [24].
 RP VARIANT CASEBROOK ASN-518.
 RX MEDLINE=91316157; PubMed=1859851;
 RA Peach R.J., Brennan S.O.;
 RT "Structural characterization of a glycoprotein variant of human serum
 albumin: albumin Casebrook (494 Asp-->Asn).";
 RL Biochim. Biophys. Acta 1097:49-54(1991).
 RN [25].
 RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
 RX MEDLINE=92190239; PubMed=1347703;
 RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
 RA Rochu D., Porta F.;
 RT "Two alloalbumins with identical electrophoretic mobility are produced


```

Query Match      100.0%; Score 3103; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.7e-194;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGGENFKALVLIAPAOYLQOCPEDEHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKSEVAHRFKDLGGENFKALVLIAPAOYLQOCPEDEHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECEFLQHKDDNPPLRLVRREV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECEFLQHKDDNPPLRLVRREV 144
QY 121 DVNCTAFHNEETFLKKYLVEIARRHPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 180
DB 145 DVNCTAFHNEETFLKKYLVEIARRHPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 204
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSQRFPKAEFAEVSCLVTDLT 240
DB 205 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSQRFPKAEFAEVSCLVTDLT 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECCKEKELLESKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECCKEKELLESKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESDVCKNYAEAKDVFLGFLYAYARRHPDYSVULLRLRAKTYETLEKC 360
DB 325 DLPSLAADFVESDVCKNYAEAKDVFLGFLYAYARRHPDYSVULLRLRAKTYETLEKC 384
QY 361 CAADAPHECYAKVDFEKFPLVEEPONLQKNCLEFQELGEYKFNALLVRYTKVPQVST 420
DB 385 CAADAPHECYAKVDFEKFPLVEEPONLQKNCLEFQELGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRLNGKYSCKCKHPEAKRWPCEADYLSVVLNQLCVLHEKTPYSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKYSCKCKHPEAKRWPCEADYLSVVLNQLCVLHEKTPYSDRVTKCCTES 504
QY 481 LVNRRPFCFSALEVDVETVVPKFEAFNAETFTPHAD:CTLSEKERQIKKOTALVELVGHKPKAT 540
DB 505 LVNRRPFCFSALEVDVETVVPKFEAFNAETFTPHAD:CTLSEKERQIKKOTALVELVGHKPKAT 564
QY 541 KEOLKAWMDFAAFVEKCCCKADDKETCFABEGKLVAAASQAALGL 585
DB 565 KEOLKAWMDFAAFVEKCCCKADDKETCFABEGKLVAAASQAALGL 609

RESULT 2
ALBU_MACMU STANDARD: PRT; 600 AA.
AC Q28522;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RW SEQUENCE FROM N.A.
RX MEDLINE=93211971; PubMed=9460152;
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA Dwyer J., Putnam F.W.
RT "cDNA and protein sequence of polymorphic macaque albumins that differ
in bilirubin binding.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413 (1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.

```

```

-!- TISSUE SPECIFICITY: Plasma.
-!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
-!- SIMILARITY: Contains 3 albumin domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: M90463; AAA36906.1;
PIR: A47391; A47391.
HSP: P02768; 187B.
InterPro: IPR000264; Serum albumin.
Pfam: PF00273; transport_prot; 3.
PRINTS: P00802; SERUMALBUMIN.
ProDom: PD002486; Serum albumin; 1.
SMART: SM00103; ALBUMIN_3.
PROSITE: PS00212; ALBUMIN; 3.
Metal-binding; Lipid-binding; Repeat; Signal; Copper.
NON TER 1 1
FT SIGNAL <1 10 BY SIMILARITY.
FT PROPEP 11 16 BY SIMILARITY.
FT CHAIN 17 600 SERUM ALBUMIN.
FT DOMAIN 17 197 ALBUMIN 1.
FT DOMAIN 204 389 ALBUMIN 2.
FT DOMAIN 396 587 ALBUMIN 3.
FT METAL 19 19 COPPER (BY SIMILARITY).
FT BINDING 256 256 BILIRUBIN (POTENTIAL).
FT DISULFID 69 78 BY SIMILARITY.
FT DISULFID 91 107 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 140 185 BY SIMILARITY.
FT DISULFID 184 193 BY SIMILARITY.
FT DISULFID 216 262 BY SIMILARITY.
FT DISULFID 261 269 BY SIMILARITY.
FT DISULFID 281 295 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match      94.8%; Score 2942; DB 1; Length 600;
Best Local Similarity 93.5%; Pred. No. 4.8e-184;
Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGGENFKALVLIAPAOYLQOCPEDEHVKLVNEVTEFAKTCVADESAAE 60
DB 17 DTHKSEVAHRFKDLGGENFKALVLIAPAOYLQOCPEDEHVKLVNEVTEFAKTCVADESAAE 76
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECEFLQHKDDNPPLRLVRREV 120
DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECEFLQHKDDNPPLRLVRREV 136
QY 121 DVNCTAFHNEETFLKKYLVEIARRHPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 180
DB 137 DVNCTAFHNEETFLKKYLVEIARRHPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 196
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSQRFPKAEFAEVSCLVTDLT 240
DB 197 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSQRFPKAEFAEVSCLVTDLT 256
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECCKEKELLESKSHCIAEVENDEMPA 300
DB 257 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECCKEKELLESKSHCIAEVENDEMPA 316

```

```
QY 301 DLPSSLAADFVSKVCKNVAEKDVLGMLFYEARHEDYVSVLLRLAKTYETTLK 360
DB 317 DLPSSLAADFVSKVCKNVAEKDVLGMLFYEARHEDYVSVLLRLAKTYETTLK 376
QY 361 CAAADPHCYAKVDFEFKFLVEEPQNLIKONCELFELQGEYKFNALLVRYTKKYPQVST 420
DB 377 CAAADPHCYAKVDFEFKFLVEEPQNLIKONCELFELQGEYKFNALLVRYTKKYPQVST 436
QY 421 PTLVEVSRNLGKVGSKCCCKHPKAKMPCADYLSVVLNGLCVLHEKTPVSDRVTKCCTES 480
DB 437 PTLVEVSRNLGKVGSKCCCKHPKAKMPCADYLSVVLNGLCVLHEKTPVSDRVTKCCTES 496
QY 481 LVNRRPCFSALEVDYVVPKFNATFTFHADICTLSEKEROIKKOTALVELVKGKPKAT 540
DB 497 LVNRRPCFSALEVDYVVPKFNATFTFHADICTLSEKEROIKKOTALVELVKGKPKAT 556
QY 541 KEQLKAVMDDFAAVFKCKCKADKCTCFABEGKGLVAASQAAL 583
DB 557 KEQLKAVMDDFAAVFKCKCKADKCTCFABEGKGLVAASQAAL 599
```

```
RESULT 3
ALBU_FELCA STANDARD; PRT; 608 AA.
ID ALBU_FELCA
AC P49064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Allergen Fel d 2).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Butheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96194824; PubMed=8647469;
RA Hilger C., Grigioni F., Kohnen M., Hentges F.;
RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
RL Gene 169:295-296(1996).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-aib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X84842; CAA59279.1; -.
CC DR PIR; JC4660; S57632.
CC DR HSPG; P02768; 1E7B.
CC DR InterPro; IPR000264; Serum albumin.
CC DR Pfam; PF00273; transport prot.; 3.
CC DR PRINTS; PR00802; SERUMALBUMIN.
CC DR ProDom; PD002486; Serum albumin; 1.
CC DR SMART; SM00103; ALBUMIN; 3.
CC DR PROSITE; PS00212; ALBUMIN; 3.
CC -----
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
KW SIGNAL 1 18
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
```

```
FT DOMAIN 404 595
FT METAL 27 27 ALBUMIN 3.
FT DISULFID 77 86 COPPER.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5F60E5F CRC64;

Query Match 84.4%; Score 2620; DB 1; Length 608;
Best Local Similarity 82.0%; Pred. No. 4e-163;
Matches 478; Conservative 52; Mismatches 53; Indels C; Gaps 0;

QY 1 DAKSEVAHRFKDLGGEENFKALVLIATAFYLAQYLQOCFFEDHVKLVNEVTEFAKTCVADESA 60
DB 25 EAHQSEIARHENDLGEHFRCLVLVAFSYLQOCFFEDHVKLVNEVTEFAKTCVADQSA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPLVRPEV 120
DB 85 NCKSLHELLGDKLCTVASLRDKYGENADCCKEPERNECFLOHKDDNPNLPLVRPEA 144
QY 121 DVMCTAHEDNEETFLKYLVEIARRHYPFYFAPELLFFAKRYKAAFTCCQAAADKAACLP 180
DB 145 DAMCTAFHENEQRLGKYLVEIARRHYPFYFAPELLFYABEYKGVFTCCCEAADAACLP 204
QY 181 KLDELDRDEGKASSAKORLKASLOKFGERAFKAVARLSORFFKAEFAEYSKLVTDLT 240
DB 205 KVDALREKVLASSAKERLKASLOKFGERAFKANSVARLSOKFFKAEFAEYSKLVTDLAK 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVDENDEMPA 300
DB 265 IHKECHGDLLECADRADLAKYICENQDSISTKLKCCGKPVLEKSHCISEVERDELPA 324
QY 301 DLPSSLAADFVSKVCKNVAEKDVLGMLFYEARHEDYVSVLLRLAKTYETTLK 360
DB 325 DLPPLAVDFVEDKCYKQYQAKDVLGTFLEYYSRRHPEYSVSLLRLAKEYEATLEK 384
QY 361 CAAADPHCYAKVDFEFKFLVEEPQNLIKONCELFELQGEYKFNALLVRYTKKYPQVST 420
DB 385 CATDDPPACYAHVDFEFKFLVEEPQNLIKONCELFELQGEYKFNALLVRYTKKYPQVST 444
QY 421 PTLVEVSRNLGKVGSKCCCKHPKAKMPCADYLSVVLNGLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCCCKHPKAKMPCADYLSVVLNGLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYVVPKFNATFTFHADICTLSEKEROIKKOTALVELVKGKPKAT 540
DB 505 LVNRRPCFSALEVDYVVPKFNATFTFHADICTLSEKEROIKKOTALVELVKGKPKAT 564
QY 541 KEQLKAVMDDFAAVFKCKCKADKCTCFABEGKGLVAASQAAL 583
DB 565 BEQLKAVMDDFGSGFVCKCAEDKEACFAEGPKLVAAQAAL 607
```

```
RESULT 4
ALBU_CANFA STANDARD; PRT; 608 AA.
ID ALBU_CANFA
AC P49822; O77705; Q9TSZ4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
```



```

Db 204 KIETMRKVLASSARORLRCSAQKFCGERALKAWSVARLSQKFKPAEFVEVTKLVNLDLTK 263
QY 241 VHTSCCHGDLLECADDADADLAKYICENQDSISSKLKCECKPILKSHGICAEVENDEMPA 300
Db 264 VKHSCCHGDLLECADDADADLAKYICNODTISISSKLKCECKPILKSHGICAEVENDEMPA 323
QY 301 DLPGLAADFFVESKDVCKQVYAEKDVFLGMFLYEVARRHPDYSVVLLRLAKTVEITLTKC 360
Db 324 NLPLTADFAEDKDVCKQVYAEKDVFLGMFLYEVARRHPDYSVVLLRLAKTVEITLTKC 383
QY 361 CAADADPHECYAKVDEPKPLVBEQNLKONCELFQOLGSEYKQFONALLVRYTKKVPQVST 420
Db 384 CAKODPHACYSTVDFKLVHLDVDEPNLKNQCDQFKLGEYGFQNALVRYTKKVPQVST 443
QY 421 PTLVEVSRNLGKVGSKCKKPEAKMPCAEYDLSVVLNCLVHLEKTPVSDRVTKKCTES 480
Db 444 PTLVEVSRNLGKVGSKCKKPEAKMPCAEYDLSVVLNCLVHLEKTPVSDRVTKKCTES 503
QY 481 LVNRRPCFSALVDETVYVPEKFAETFTFHADICTLSEKERQIKKQTAALVELVKKPKAT 540
Db 504 LVNRRPCFSALVDETVYVPEKFAETFTFHADICTLSEKERQIKKQTAALVELVKKPKAT 563
QY 541 KEQLKAVMDPFAAFVEKCKKADKDETCFAEKGKLVAAASQAL 583
Db 564 BEQLKTYMENFVAFVDCCKCAADKDETCFAEKGKLVVSTQAL 606

RESULT 7
ALBU_RABIT STANDARD; PRT; 608 AA.
AC P49065;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Liver;
RA Sheffield W.P., Syed S., Schuyler P.D.;
RL Submitted (DRC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18344; AAB58347.1; -.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.

```

```

FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 68914 MW; CF5E92647AAPE9A2 CRC64;

Query Match 78.8%; Score 2446; DB 1; Length 608;
Best Local Similarity 74.1%; Pred. No. 8.1e-152;
Matches 433; Conservative 77; Mismatches 74; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKOLGEENFKALVLIAPAYLOQCPEDHDVKNLNEVTEPAKTCAVDESAE 60
DB 25 EAHKSEIAHRFNDVGEHFGLVLITFSQYLOKCPYEEHAKLVKEVTDLAKACVADSA 84
QY 61 NCDKSLHTLFGDKLCTVATREYTGEMACCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 85 NCDKSLHDFGDKLCTVATREYTGEMACCAKQEPERNECFLOHKDDNPRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKKLYEARRHPYFYAPYELFFAKRYKAAFTCCQAAADKACLLP 180
DB 145 DVLCKAFHDDEKAFHGHLYEVARRHPYFYAPYELLYYAKYKALTECCCEAADKACLLP 294
QY 181 KLDELDEGKASSAKQRLKCAKQGERAFKAWAVARLSORFPKAEPAEVSCLVTLDTK 240
DB 205 KLDALEKGLISAAQERLRCASIKQFDRAYKAWALVRLSQRFPKADFTDISKIVTLTK 264
QY 241 VHTSCCHGDLLECADDADADLAKYICENQDSISSKLKCECKPILKSHGICAEVENDEMPA 300
DB 265 VKHSCCHGDLLECADDADADLAKYICENQDSISSKLKCECKPILKSHGICAEVENDEMPA 324
QY 301 DLPGLAADFFVESKDVCKQVYAEKDVFLGMFLYEVARRHPDYSVVLLRLAKTVEITLTKC 360
DB 325 GLPAVAEEFVEDKDVCKQVYAEKDVFLGMFLYEVARRHPDYSVVLLRLAKTVEITLTKC 384
QY 361 CAADADPHECYAKVDEPKPLVBEQNLKONCELFQOLGSEYKQFONALLVRYTKKVPQVST 420
DB 385 CATDDPHACYAKVDEPKPLVBEQNLKONCELFQOLGSEYKQFONALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKKPEAKMPCAEYDLSVVLNCLVHLEKTPVSDRVTKKCTES 480
DB 445 PTLVEVSRNLGKVGSKCKKPEAKMPCAEYDLSVVLNCLVHLEKTPVSDRVTKKCTES 504
QY 481 LVNRRPCFSALVDETVYVPEKFAETFTFHADICTLSEKERQIKKQTAALVELVKKPKAT 540
DB 505 LVNRRPCFSALVDETVYVPEKFAETFTFHADICTLSEKERQIKKQTAALVELVKKPKAT 564
QY 541 KEQLKAVMDPFAAFVEKCKKADKDETCFAEKGKLVAAASQALG 584
DB 565 NQJKTUVGEFTALLDKCCSAEDKEACFAVEGPKLVESKATLG 608

RESULT 8
ALBU_SHEEP STANDARD; PRT; 607 AA.
ID ALBU_SHEEP
AC P14639;

```


"Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage of rat serum albumin.";
 J. Biochem. 83:35-48(1978).
 [4]
 SEQUENCE OF 223-288 AND 572-608.
 MEDLINE=76260153; PubMed=956149;
 Isemura S., Ikenaka T.;
 "Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino acid sequences of four fragments.";
 J. Biochem. 79:1183-1196(1976).
 [5]
 SEQUENCE OF 166-174.
 TISSUE=Plasma;
 MEDLINE=87194805; PubMed=2437111;
 Carraway R.E., Mitra S.P., Cochrane D.E.;
 "Structure of a biologically active neurotensin-related peptide obtained from pepsin-treated albumin(s).";
 J. Biol. Chem. 262:5968-5973(1987).
 [6]
 COPPER-BINDING.
 MEDLINE=79001617; PubMed=80265;
 Aoyagi Y., Ikenaka T., Ichida F.;
 "Copper(II)-binding ability of human alpha-fetoprotein.";
 Copper Res. 38:3483-3486(1978).
 -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.
 -!- FUNCTION: NRP REGULATES FAT DIGESTION, LIPID ABSORPTION, AND BLOOD FLOW (POTENTIAL).
 -!- SUBCELLULAR LOCATION: Secreted.
 -!- TISSUE SPECIFICITY: Plasma.
 -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 -!- SIMILARITY: Contains 3 albumin domains.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; V01222; CAA24532.1; .
 PIR; A93872; ABRTS.
 DR HSSP; P02768; 1E7B.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608 SERUM ALBUMIN.
 FT PEPTIDE 166 174 NEUTROTENSIN-RELATED PEPTIDE.
 FT DOMAIN 25 205 ALBUMIN 1.
 FT DOMAIN 212 397 ALBUMIN 2.
 FT DOMAIN 404 595 ALBUMIN 3.
 FT METAL 27 27 COPPER.
 FT METAL 77 86 BY SIMILARITY.
 FT DISULFID 99 115 BY SIMILARITY.
 FT DISULFID 114 125 BY SIMILARITY.
 FT DISULFID 148 193 BY SIMILARITY.
 FT DISULFID 192 201 BY SIMILARITY.
 FT DISULFID 224 270 BY SIMILARITY.
 FT DISULFID 269 277 BY SIMILARITY.
 FT DISULFID 289 303 BY SIMILARITY.
 FT DISULFID 302 313 BY SIMILARITY.
 FT DISULFID 340 385 BY SIMILARITY.
 FT DISULFID 384 393 BY SIMILARITY.
 FT DISULFID 416 462 BY SIMILARITY.

FT DISULFID 461 472 BY SIMILARITY.
 FT DISULFID 485 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT VARIANT 262 262 V -> L (IN REF. 5).
 FT CONFLICT 174 174 Y -> L (IN REF. 5).
 SQ SEQUENCE 608 AA; 68718 MW; 5B8497A282411AB7 CRC64;
 Query Match 78.2%; Score 2426; DB 1; Length 608;
 Best Local Similarity 73.4%; Pred. No. 1.6e-150;
 Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;
 Oy 1 CAHKSEVAHRFDKLGEEFKALVLIAPAOYLOOCFFEDHVKLVNVEVTEFAKTCVADSEAE 60
 Db 25 EAHKSEIAHRFDKLGEOHFGLVLIAPSOYLOKCPYEENIKLVQEVTFDAKTCVADENAE 84
 Oy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPRLVLRPEV 120
 Db 85 NCDKSIHTLFGDKLCAIPKLRDNYGELADCCAKQPERNECFLOHKDDNPRLVLRPEA 144
 Oy 121 DVNCTAFHNDNEETFLKKYLYEIAARRHPYFAPDELLFFAKRYKAAFTCCQADAAACLLP 180
 Db 145 EAMCTSFQENPTSFGLHYLHEVARRHYPYFAPDELLYYAEKYNEVLTQCTESDKAACLTP 204
 Oy 181 KLDELDECKKASSAKORLAKCASIQKFGEPAPKAMAVARLSORFPKAEFVSKLVTDLTK 240
 Db 205 KLDVAREKALVAVRQRMKSSMORPGERAFKAMAVARMSORFPNAEFAEITKATDVTK 264
 Oy 241 VHTCCCHGDLLECADRADLAKYICENQDSISKLKECCPKLLEKSHCAEVENDEMPA 300
 Db 265 INKECCHGDLLECADRADLAKYMCENQATISSKLQACCDKPVQKSOCLAETEDHNIPA 324
 Oy 301 DLPSLAADPVESKOVCKNYAEAKDVLGMLFYEARHPDYSVLLLLAKYETITLKC 360
 Db 325 DLPSIAADPVEDKEVCNVAEAKDVLGTLFYESRRHPDYSVLLLLAKYETITLKC 384
 Oy 361 CAAADPHCYAKVDFPKP-VSEPNLIKONCELPFOELGEYKFONALVRYTKKYPQVST 420
 Db 385 CAGDPPACVGTVLAEFQPLVEEPKMLVKTNCELYEKUFGFQNAVLYRTCKAPQVST 444
 Oy 421 PTLVFSRNLGKVGSKCKKHPEAKRMPCAEYLSVVLNQLCVLHEKTPVSDRTKCTES 480
 Db 445 PTLVEAARNLGRVGTCKC-LPBAQRPCVEDYLSAILNRLCVLHEKTPVSEKVTCKCSG 504
 Oy 481 LVNRRPCFSALEVDETYVPKENAEFTTFAHDCITLSEKEROIKKOTALVELVKKPKAT 540
 Db 505 LVERRPCFSALTVDETYVPKEFAEFTTFSIDICITLPEKEROIKKOTALVELVKKPKAT 564
 Oy 541 KEQLKAVMDGFAAFVEKCKCKADKDKETCFAEEGKLVAAASQAAL 583
 Db 565 EDQLKTVMGDFAQFVDKCKCAADKDKNCFCATEGNLVARSKEAL 607
 RESULT 10
 ID ALBU_PIG STANDARD; PRT: 605 AA.
 AC P08835; Q29018;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor (fragment).
 GN ALB.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9923;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RC MEDLINE=89016582; PubMed=3174440;
 RA Baldwin G.S., Weinstein J.;
 RT "Nucleotide sequence of porcine liver albumin.";

RL Nucleic Acids Res. 16:9045-9045(1988).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Plasma.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X12422; CAA30970.1; -.
 DR EMBL; M36787; AAA30988.1; -.
 DR PIR; SGL382; ABPGS.
 DR HSP; P02768; IE7B.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN_3.
 DR PROSITE; PS00212; ALBUMIN_3.
 DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 KW METAL-BINDING; LIPID-BINDING; REPEAT; SIGNAL; COPPER.
 FT NON_TER 1 1
 FT SIGNAL <1 16 BY SIMILARITY.
 FT PROPEP 17 22 BY SIMILARITY.
 FT CHAIN 23 605 SERUM ALBUMIN.
 FT DOMAIN 23 202 ALBUMIN 1.
 FT DOMAIN 209 394 ALBUMIN 2.
 FT DOMAIN 401 592 ALBUMIN 3.
 FT METAL 31 31 COPPER (BY SIMILARITY).
 FT METAL 75 84 BY SIMILARITY.
 FT DISULFID 97 113 BY SIMILARITY.
 FT DISULFID 112 123 BY SIMILARITY.
 FT DISULFID 145 193 BY SIMILARITY.
 FT DISULFID 189 198 BY SIMILARITY.
 FT DISULFID 221 267 BY SIMILARITY.
 FT DISULFID 266 274 BY SIMILARITY.
 FT DISULFID 286 303 BY SIMILARITY.
 FT DISULFID 299 310 BY SIMILARITY.
 FT DISULFID 337 382 BY SIMILARITY.
 FT DISULFID 381 390 BY SIMILARITY.
 FT DISULFID 413 459 BY SIMILARITY.
 FT DISULFID 458 469 BY SIMILARITY.
 FT DISULFID 482 498 BY SIMILARITY.
 FT DISULFID 497 508 BY SIMILARITY.
 FT DISULFID 535 580 BY SIMILARITY.
 FT DISULFID 579 588 BY SIMILARITY.
 FT DISULFID 562 562 E -> D (IN REF. 1; AAA30988).
 FT CONFLICT 562 562
 FT SEQUENCE 605 AA; 694-0 MW; 3E556B0DD1A1F4FF CRC64;
 SQ
 Query Match 77.7%; Score 241.1.5; DB 1; Length 605;
 Best Local Similarity 76.0%; Pred. No. 1.4e-149;
 Matches 438; Conservative 67; Mismatches 70; Indels 1; Gaps 1;
 1 DAHSEVAHFRKDLGGEENFKALVIAFYLOQCPFEHVKLVNEVTFBKTCTVADESAAE 60
 23 DTYKSEIAHFRKDLGEOYFKGLVLAFLPSHLOQCPYEEHVKLVRETVTFBKTCTVADESAAE 82
 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRILRVREV 120
 83 NCDKSLHTLFGDKLCAIPSLREHYGDLADCCKEPERNECFLOHKDNPDIPLKL-KDPP 141
 121 DWMTAFHNEETFLKKYLYEARRHPYAPPELLFFFAKRYKAAFTCCQAADKAAACLLP 180
 142 VALCADFOEDQKFMGKYLIEARRHPYAPPELLYYAIYYKDVFSQCQADKAAACLLP 201

181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
 202 KIEHLREKVLTSAAKORLKASIQKFGERAFAKAWAVARLSORFPKADFTIEISKIVTDLAK 261
 241 VHTCCCHGDLLECADDRAADLAKYICENQDSISSKLKECCCEKPLEKSHCIAEVENDEMPA 300
 262 VHKECCCHGDLLECADDRAADLAKYICENQDTISTKLKECCCDKPLEKSHCIAEAKRDELPA 321
 301 DLPSLAADFVESKDVCKNYAEAKDVGFMFLFYEYARRHDPDYSVLLRLAKYETITLEKC 360
 322 DLNPLHDFVEDKEVCNKYAEAKDVGFLGTFLFYSRRHDPDYSVLLRLAKYETITLEKC 381
 361 CAAADPEHCYAKVDFEPKPLVREPQNLIKONCELPOLGEYKFNALLVRYTKKVPQVST 420
 382 CAXEDPPACYATVDFEPKPLVREPQNLIKONCELPOLGEYKFNALLVRYTKKVPQVST 441
 421 PTLVPSVRLGKVGSKCKHPKAPKAEADYLSVNLQCVLHEKTPVSDRVTKCTES 480
 442 PTLVEVARKLGLVGRCKCRPEERLSCAEDYLSVNLQCVLHEKTPVSEKVTCKCTES 501
 481 LVNRRCFSALEVDYTVPKFNAETFTPHADICTLSEKEROIKKOTALVELVKKHKKPAT 540
 502 LVNRRCFSALEVDYTVPKFNAETFTPHADICTLSEKEROIKKOTALVELVKKHKKPAT 561
 541 KEOLKAVMDDPFAAFVEKCKCKADDKETCFPAEKGKLV 576
 562 EEQLRTVLGNFAAFVQKCAAPDHEACPAVEGPKFV 597
 RESULT 11
 ID ALBU_MERUN STANDARD; PRT; 609 AA.
 AC Q35030;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
 OC Meriones.
 RX NCBI_TaxID=10047;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MCS IDR; TISSUE=Liver.
 RX MEDLINE=98116663; PubMed=9455485;
 RA Yoshida K., Seto-Onshima A., Sinohara H.;
 RT "Sequencing of cDNA encoding serum albumin and its extrahepatic
 RL synthesis in the Mongolian gerbil, Meriones unguiculatus.";
 RL DNA Res. 4:351-354(1997).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Plasma.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB006197; BAA21765.1; -.
 DR PIR; JCS838; JCS838.
 DR HSP; P02768; IE7B.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.

```

DR PRINTS; PRO0802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 609
FT DOMAIN 25 206
FT DOMAIN 213 398
FT DOMAIN 405 596
FT METAL 28 28
FT DISULFID 78 87
FT DISULFID 100 116
FT DISULFID 115 126
FT DISULFID 149 194
FT DISULFID 193 202
FT DISULFID 225 271
FT DISULFID 270 278
FT DISULFID 290 304
FT DISULFID 303 314
FT DISULFID 341 386
FT DISULFID 385 394
FT DISULFID 417 463
FT DISULFID 462 473
FT DISULFID 486 502
FT DISULFID 501 512
FT DISULFID 539 584
FT DISULFID 583 592
SQ SEQUENCE 609 AA; 68940 MW; 9CA5F97F67BF1A48 CRC64;

Query Match
Best Local Similarity 73.98; Score 2387; DB 1; Length 609;
Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;

Qy 2 AHKSEVAHRPKDGEENFKALVLIAPQYQQCFEDHVKLVNEVTEFAKTCVADSAEN 61
Db 27 AHKSEIAHRYKDLGEKFKGLVLYTFSGYKQKCSYEERVKVREVTDFASNCADSAEN 96
Qy 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKOPEPERNECFLOHKDNNPLPLVPEVD 121
Db 87 CDKSLHTLFGDKLCSLPNFKGEYKAEADCCAKOPEPERNECFLOHKDNNPLPLVPEVD 146
Qy 122 VWCATFADHNEETFKKLYLEIARRHPFYAPPELLFFAKRYKAAFTCCOAAKAAALLPK 181
Db 147 AMCTAFQENAEAFMGHYLHEVARRHPFYFGPELLYLADKYTAVLTCCAADKAGCLTPK 206
Qy 182 LDEURDEGKASAKQRLKCAQLQFGBRFAKAMAVARLSORFPAKFAEVSKLVTLTKV 241
Db 207 LDALKEKALVSARQRLKCSMKKFGGERAFKAMAVARMSQTFNADFAEITKLATLTKV 266
Qy 242 HTECHGDLECADRADLAKYICENODSLSKKLKECCKEKPLLEKSHCIAEVENDEWPD 301
Db 267 TQECCHGDLECADRADLAKYICENODSLSKKLKECCKEKPLLEKSHCIAEVENDEWPD 326
Qy 302 LPSLAADPVEKSKYCKNYAEAKDYFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLKQC 361
Db 327 LPALTADPVEDKDYCKNYAEAKDYFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLKQC 386
Qy 362 AAADPHCYAKVDFEFPFLVPEEPONLQKNCLEPEQGEYKFNALLVRYTKVPOVSTP 421
Db 387 AEADPHCYGHVDFEFPFLVPEEPONLQKNCLEPEQGEYKFNALLVRYTKVPOVSTP 446
Qy 422 TLVEVSRNLGKVGKCKHPKAKMPCAEADYLSVVLNOLCVLHEKTPVSDRVTKCCTESL 481
Db 447 TLVEAARSLGRVTHCCALPEKKRLPCVEDYLSAILNRVCLLHEKTPVSDRVTKCCTESL 506
Qy 482 VNRPCFSALEVDETYVPKFNASTFTFHADICTLSEKQRQIKQTALVELVHKPKAAK 541
Db 507 VERRPCFSALPVDETYVPKFKASTFTFHANICTLPEKQKQKQALVELVHKPKQATE 566
Qy 542 EQLKAVMDFFAAFEVKCKKADDKETCTFAEKGKLVAAEQAL 583

```

```

Db 567 EQLKVMGDFAEFLKCKQEDKEACFSTEGPKLVAESQAL 608

RESULT 12
ALBU_MOUSE
ID ALBU_MOUSE STANDARD; PRT; 608 AA.
AC P07724; 061802;
DT 01-APR-1988 (Rel. 07, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB OR ALB1 OR ALB-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC TISSUE=Liver;
RP SEQUENCE FROM N.A.
RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saio T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Satalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml J.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Rono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fietcher C., Fujita M., Gariboldi M.,
RA Guskinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez J., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [3]
RP Nature 409:685-690(2001).
RX SEQUENCE OF 99-516 FROM N.A.
RX MEDLINE=88216123; PubMed=2452956;
RA Mirghetti P.P., Law S.W., Dugaiczky A.;
RT "The rate of molecular evolution of alpha-fetoprotein approaches that
RT of pseudogenes.";
RL Mol. Biol. Evol. 2:347-358(1985).
RN [4]
RP SEQUENCE OF 477-551 FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=90269606; PubMed=1971802;
RA Boccaccio C., Deschattratte J., Meunier-Rotival M.;
RT "Empty and occupied insertion site of the truncated LINE-1 repeat
RT located in the mouse serum albumin-encoding gene.";
RL Gene 88:181-186(1990).
RN [5]
RP SEQUENCE OF 25-44.
RC TISSUE=Liver;
RX MEDLINE=93162044; PubMed=1286668;
RA Giometti C.S., Taylor J., Tollaksen S.L.;
RT "Mouse liver protein database: a catalog of proteins detected by two-
RT dimensional gel electrophoresis.";
RL Electrophoresis 13:970-991(1992).
CC !- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC !- SUBCELLULAR LOCATION: Secreted.

```


CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X60688; CAA43098.1; -
 DR EMBL: V00381; CAA23680.1; -
 DR PIR: S15571; ABCHS.
 DR HSP: P02768; IEB.
 DR InterPro: IPR000264; Serum albumin.
 DR Pfam: PF00273; transport prot.; 3.
 DR PRINTS: PR00802; SRUMALBUMIN.
 DR ProDom: PD002486; Serum albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 23
 FT CHAIN 24 615 SERUM ALBUMIN.
 FT DOMAIN 24 209 ALBUMIN 1.
 FT DOMAIN 216 401 ALBUMIN 2.
 FT DOMAIN 408 599 ALBUMIN 3.
 FT METAL 30 30 COPPER (BY SIMILARITY).
 FT DISULFID 80 89 BY SIMILARITY.
 FT DISULFID 102 118 BY SIMILARITY.
 FT DISULFID 117 128 BY SIMILARITY.
 FT DISULFID 152 197 BY SIMILARITY.
 FT DISULFID 196 205 BY SIMILARITY.
 FT DISULFID 228 274 BY SIMILARITY.
 FT DISULFID 273 281 BY SIMILARITY.
 FT DISULFID 293 307 BY SIMILARITY.
 FT DISULFID 306 317 BY SIMILARITY.
 FT DISULFID 344 389 BY SIMILARITY.
 FT DISULFID 388 397 BY SIMILARITY.
 FT DISULFID 420 466 BY SIMILARITY.
 FT DISULFID 465 476 BY SIMILARITY.
 FT DISULFID 489 505 BY SIMILARITY.
 FT DISULFID 504 515 BY SIMILARITY.
 FT DISULFID 542 587 BY SIMILARITY.
 FT DISULFID 586 595 BY SIMILARITY.
 FT CARBOHYD 500 500 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 24 24 F -> M (IN REF. 3).
 SQ SEQUENCE 615 AA; 69918 MW; B59E48BCAEC066C6 CRC64;
 Query Match 50.2%; Score 1557.5; DB 1; Length 615;
 Best Local Similarity 46.7%; Pred. No. 4.3e-94;
 Matches 273; Conservative 18; Mismatches 192; Indels 1; Gaps 1;
 QY 3 HKSEVAHREKDLGEEFKALVLIQYLCQPPEDHVKLVNEVTEFAKTCVADESAENC 62
 DB 30 HKSEIAHRYNDLKEETFKAVAMITFAQLQRCSEGLSKLVKQVDLQACKVANEAPEC 89
 QY 63 DKSHTLFGDKLCTVATLRETYGEMACCAQEPERNECFLOHKDDNPLR-LVREPEVD 121
 DB 90 SKLPISILDEICQVEKLRDSYGAMVCCSKAOPERNECFLSFKVSPQDFVQYQRPASD 149
 QY 122 VMCTAFHDNEETFLKLVLYELARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAGLLPK 181
 DB 150 VICQEQDNRVSLFGHFIYSVARRHPFLYAPAILSFVDFEHALQSCCKESDVGACLDYK 209
 QY 182 LDELURDEKASAKORUKCASLQKGERAFKAMAVARLSQRPFAKFAEVSKLVTDLTKV 241
 DB 210 EIVMREKAGSVKQYVFCGILKQGRVFOARQLIYLSQKYPKAPFSEVSKFVHDSIGV 269
 QY 242 HTECHGDLLECADRADLAKYICENQDSTSSKLEKCEKPLEKSKCIAEVENDEMPAD 301
 DB 270 HKECEGDMVECMDDAMMNSLCSQDDVFSGKTKDCCEPIVERSQICINAEFDEKPAD 329
 QY 302 LPSLAADPVEKDKCKYAEAKDVFLGFLYELARRHPDYSVLLRLRLKATYETILEKCC 361
 DB 330 LPSLVKDYIEDKVECKSPFAGHDFAFMAFVVEYSRRHPFSQLIMRIAKGYESLLEKCC 389

QY 362 AADPHCEYAKVDFEFKPLVEBPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVSTP 421
 DB 390 KTONPABCYANAEQUNQHOKETQDVVVKTNCDLJHDHGEADFLKSLIRYKKMPQVPTD 449
 QY 422 TLVEVSNLGKVSCKCKHPKAPKMPKCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTESL 481
 DB 450 LLLGTGKKMTTIGTKCQQLGEDRRMACSEGYLSIVIHDTCKRQETTPINDNVSQCSQLY 509
 QY 482 VNRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQIALVELVGHKPKATK 541
 DB 510 ANRRPCFTAMGVDTKYVPPFPNPFDMFSEKLCSPAEEBEVGMKLLINLIRKPKQMT 569
 QY 542 EQLKAVMDDEFAAFVEKCKKADDEKTCFAERGGKLVAAASQAALGL 585
 DB 570 EQIKTIADGETAMVDKCKCKQSDINTCEGEGANLIVQSRATLGI 613
 RESULT 14
 FETA_PANTR
 ID FETA_PANTR STANDARD; PRT; 609 AA.
 AC Q28789;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).
 DE fetoprotein).
 GN AFP.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96032345; PubMed=7557431;
 RA Nishio H., Gibbs P.E., Minghetti P.P., Zielinski R., Dugaiczak A.;
 RT "The chimpanzee alpha-fetoprotein-encoding gene shows structural
 RT similarity to that of gorilla but distinct differences from that of
 RT human.";
 RL Gene 162:213-220(1995).
 CC -!- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND
 CC BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE
 CC (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.
 CC -!- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION
 CC TO THE MONOMERIC FORM.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and
 CC yolk sac.
 CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U21916; AAA91641.1; -
 DR PIR: JC4258; JC4258.
 DR HSP: P02768; IEB.
 DR InterPro: IPR000264; Serum albumin.
 DR Pfam: PF00273; transport prot.; 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 2.
 KW Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;
 FT SIGNAL 1 18 BY SIMILARITY.
 FT CHAIN 19 609 ALPHA-FETOPROTEIN.
 FT DOMAIN 20 205 ALBUMIN 1.
 FT DOMAIN 212 397 ALBUMIN 2.

```

FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 22 22 COPPER AND NICKEL (BY SIMILARITY).
FT DISULFID 99 114 BY SIMILARITY.
FT DISULFID 113 124 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 609 AA; 68741 VM; C032987CAD0E672B CRC64;

Query Match 40.4%; Score 1253.5; DB 1; Length 609;
Best Local Similarity 40.1%; Pred. No. 2.4e-74;
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;

QY 3 HKSE-----VAHRFKDLGLENFKALVLIFAFAYLOQCPEFDEHVKLVNEVTFPAKTCVADE 57
DB 22 HNEVGIASILSYOCTAENITLDTATFFAQVQEAITYKEVSKVXDALTAIEKPTGDE 81
QY 58 SAENCOKSHTT.FGDKLCTVATLRETYGEMACCAKQBPENECFLOHKDNP.NLPLRV 116
DB 82 QSAGCLENLPAFLBELCREKEILEKYGH.SDCCSQSBEGRNCFLAHKKTPPASIPFPFQ 140
QY 117 REVDVMTAFHNDSETLKKYLYETARHPHYEYAPELLFFAKRYKAAFTCCQADKAA 176
DB 141 VPEPTSCAYBEDRETFMKNKIYETARHPHYEYAPELLFFAKRYKAAFTCCQADKAA 200
QY 177 CLLPKLDELREDEGKASSAKQRLKASLOKFGBRFAKMAVARLSQFPAEVSKLVT 236
DB 201 CFOTKAATVTKELRESSLLNQHACAVMKNGFRTFOAITVTKLSQKFTKWNFTLOKLVL 260
QY 237 DUTKVHTECHDGLLECAADRADLAKYICENODSISSKLKECEKPLLEKSHCIAEVND 296
DB 261 DVARVHEHCRCGV.DCLQDGEKIMSYICSQDITLSNKITECKLTLTLEGGCIIHAEND 320
QY 297 ENPADLPSLAADPVESKVKYAEAKQVFLGMELYEYARRHPDYVSVLLRLAKTYETT 356
DB 321 EKPEGLSPN.NRFLGDRDFNQFSSGEKNIFLASFVHEYSRRHFQAVSVILRVAKGYDEL 380
QY 357 LEKCCAAADPHCYAKVDFEPLVPEEPQNLKQNCLEFQELGEYKFFONALLVRYTKKVP 416
DB 381 JEKCFOTENPLECQDGEELQYIQESQALAKRSCGLFQKLGVEYVLQNAFLVATKAP 440
QY 417 QVSTFTLVSRNLGVKSGKCKHPKAMPKCAEDYLSVLNQLCVLHEKTPVSRVTKC 476
DB 441 QLTSELMAITRAATATAATCCQLEDKLLACGEAARD:IIHGLCIRHETTPNPNVGVQC 500
QY 477 CTESLWNRPCPSALEVDETYVPKFPNFAETFTFHADICTLSEKERQIKKQATLALVELVKH 536
DB 501 CTSVYANRRPCFSSLVVDFTYVPPAFSDSKFTHDKDLCQAQGVALQTKNQEFINLVKX 560
QY 537 PKATKEQLXAVMDDFAAFPVEKCKADKDETCFAEKGKLLVAASQAALGL 585
DB 561 PQITEQLEAVIADFGLLEKCCQGEQVCFVCEGQKLSIKTRAALGV 609

RESULT 15
FETA HUMAN
ID FETA HUMAN STANDARD; PRT; 609 AA.
AC P0271;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-f-
DE fetoprotein).
DE APP.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83273664; PubMed=6192439;
RA Morinaga T., Sakai M., Wegmann T.G., Tamaoki T.;
RT "Primary structures of human alpha-fetoprotein and its mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4604-4608(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87185438; PubMed=2436661;
RA Gibbs P.E.M., Zieleniski R., Boyd C., Dugaiczak A.;
RT "Structure, polymorphism, and novel repeated DNA elements revealed by
a complete sequence of the human alpha-fetoprotein gene.";
RL Biochemistry 26:1332-1343(1987).
RN [3]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=93278385; PubMed=7684942;
RA McVey J.H., Michaelides K., Hansen L.P., Ferguson-Smith M.,
RA Tilghman S., Krumlauf R., Tuddenham E.G.D.;
RT "A G-->A substitution in an HNF I binding site in the human alpha-
fetoprotein gene is associated with hereditary persistence of alpha-
fetoprotein (HPAFP).";
RL Hum. Mol. Genet. 2:379-379(1993).
RN [4]
RP SEQUENCE OF 429-556 FROM N.A.
RX MEDLINE=83158778; PubMed=6187626;
RA Beattie W.G., Dugaiczak A.;
RT "Structure and evolution of human alpha-fetoprotein deduced from
partial sequence of cloned cDNA.";
RL Gene 20:415-422(1982).
RN [5]
RP PARTIAL SEQUENCE OF 19-609.
RX MEDLINE=91242409; PubMed=1709810;
RA Pucci P., Scigliano R., Malorni A., Marino G., Tecce M.F.,
RA Ceccarini C., Ferrana B.;
RT "Human alpha-fetoprotein primary structure: a mass spectrometric
study.";
RL Biochemistry 30:5061-5066(1991).
RN [6]
RP PRELIMINARY SEQUENCE OF 19-35.
RX MEDLINE=77242506; PubMed=70228;
RA Yachnin S., Hsu R., Heinrichson R.L., Miller J.B.;
RT "Studies on human alpha-fetoprotein. Isolation and characterization
of monomeric and polymeric forms and amino-terminal sequence
analysis.";
RL Biochim. Biophys. Acta 493:418-428(1977).
RN [7]
RP PRELIMINARY SEQUENCE OF 19-38.
RX MEDLINE=78001760; PubMed=71198;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Comparative chemical structures of human alpha-fetoproteins from
fetal serum and from ascites fluid of a patient with hepatoma.";
RL Cancer Res. 37:3663-3667(1977).
RN [8]
RP PRELIMINARY SEQUENCE OF 19-39.
RX MEDLINE=75018719; PubMed=4138095;
RA Ruoslahti E., Pihko H., Vaheri A., Seppala M., Virolainen M.,
RA Kontinen A.;
RT "Alpha fetoprotein: structure and expression in man and inbred mouse
strains under normal conditions and liver injury.";
RL Johns Hopkins Med. J. Suppl. 3:249-255(1974).
RN [9]
RP GENE STRUCTURE.
RX MEDLINE=85182629; PubMed=2580830;
RA Sakai M., Morinaga T., Urano Y., Watanabe K., Wegmann T.G.,
RA Tamaoki T.;
RT "The human alpha-fetoprotein gene. Sequence organization and the 5'

```

RT flanking region.";
 RL J. Biol. Chem. 260:5055-5060(1985).
 RN [10]
 RP METAL-BINDING.
 RX MEDLINE=79001617; PubMed=80265;
 RA Aoyagi Y., Ikenaka T., Ichida F.;
 RT "Copper(II)-binding ability of human alpha-fetoprotein.";
 RL Cancer Res. 38:3483-3486(1978).
 RN [11]
 RP BILIRUBIN-BINDING.
 RX MEDLINE=80001710; PubMed=89500;
 RA Aoyagi Y., Ikenaka T., Ichida F.;
 RT "Alpha-fetoprotein as a carrier protein in plasma and its bilirubin-binding ability.";
 RL Cancer Res. 39:3571-3574(1979).
 RN [12]
 RP SULFATION.
 RX MEDLINE=86042625; PubMed=2414772;
 RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;
 RT "Tyrosine sulfation of proteins from the human hepatoma cell line HepG2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).
 CC -!- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.
 CC -!- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION TO THE MONOMERIC FORM.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: PLASMA. SYNTHESIZED BY THE FETAL LIVER AND YOLK SAC.
 CC -!- DEVELOPMENTAL STAGE: OCCURS IN THE PLASMA OF FETUSES MORE THAN 4 WEEKS OLD, REACHES THE HIGHEST LEVELS DURING THE 12TH-16TH WEEK OF GESTATION, AND DROPS TO TRACE AMOUNTS AFTER BIRTH. THE SERUM LEVEL IN ADULTS IS USUALLY LESS THAN 40 NG/ML. AFP OCCURS ALSO AT HIGH LEVELS IN THE PLASMA AND ASCITIC FLUID OF ADULTS WITH HEPATOMA.
 CC -!- PTM: INDEPENDENT STUDIES SUGGEST HETEROGENEITY OF THE AMINO-TERMINAL SEQUENCE OF THE MATURE PROTEIN AND OF THE CLEAVAGE SITE OF THE SIGNAL SEQUENCE.
 CC -!- PTM: SULFATED.
 CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M10949; AAA51674.1; -
 CC EMBL; M10950; AAA51675.1; -
 CC EMBL; V01514; CAA24758.1; -
 CC EMBL; M16110; BAB58754.1; -
 CC EMBL; Z19532; CAA79592.1; -
 CC PIR; A26624; FPHU.
 CC HSSP; P02768; IE7B.
 CC GlycoSuiteDB; P02771; -
 CC Siena-2DPAGE; P02771; -
 CC Genew; HGNC:317; AFP.
 CC MIM; 104150; -
 CC InterPro; IPR000264; Serum_albumin.
 CC Pfam; PF00273; transport_prot; 3.
 CC PRINTS; PR00802; SERUMALBUMIN.
 CC ProDom; PD002486; Serum_albumin; 1.
 CC SMART; SM00103; ALBUMIN; 3.
 CC PROSITE; PS00212; ALBUMIN; 2.
 CC Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;
 KW Signal; Polymorphism.
 FT SIGNAL 1 18
 FT CHAIN 19 609 ALPHA-FETOPROTEIN.
 FT DOMAIN 20 205 ALBUMIN 1.
 FT DOMAIN 212 397 ALBUMIN 2.

FT	DOMAIN	404	595	ALBUMIN 3.
FT	METAL	22	22	COPPER AND NICKEL.
FT	DISULFID	99	114	
FT	DISULFID	113	124	
FT	DISULFID	148	193	
FT	DISULFID	192	201	
FT	DISULFID	224	270	
FT	DISULFID	269	277	
FT	DISULFID	289	303	
FT	DISULFID	302	313	
FT	DISULFID	384	393	
FT	DISULFID	416	462	
FT	DISULFID	461	472	
FT	DISULFID	485	501	
FT	DISULFID	500	511	
FT	DISULFID	538	583	
FT	DISULFID	582	591	
FT	CARBOHYD	251	251	
FT	VARIANT	570	570	N-LINKED (GLCNAC. . .). /FTID=CAR 000070. A -> G (in dbSNP:1057173). /FTID=VAR 012049.
FT	SEQUENCE	609 AA;	68677 MM;	4D4E45820E1C2D4F CRC64;
QY	Query Match	40.3%;	Score 1249.5;	DB 1; Length 609;
Db	Best Local Similarity	39.9%;	Pred. No. 4.4e-74;	
QY	Matches 235;	Conservative 116;	Mismatches 231;	Indels 7; Gaps 3;
QY	3 HKSE----	VAHRFKDLGENFKALVLIAPAOYLQCPFEDHVKLVNEVTEFAKTCVADE	57	
Db	22 HRNEYGIASLDSYQCTAEISLADLATIFPAQVQEAITYKEVSKMKDALTAEKPTGDE	81		
QY	58 SAENDCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEKNECFLOHKDDNP-NLPLRV	116		
Db	82 QSSGCLNQPLFALEELCHEKEILEKYGH-SDCCSQSEGRHNCFLAHKKPTPASIPLFQ	140		
QY	117 RPEVDVMTAFHNEETFLKKYLYETARRHPYFAPPELLFFAKRYKAAATECCQAAKAA	176		
Db	141 VPEPVTSCAEYEDRETFMNKFYEIARRHPPIFYAFTILLWAARYDKIIPSCCKNAVE	200		
QY	177 CLLPKLDELDEGKASSAKORLKCASLQKFGFRKAWAVARLSQRPFAEFAEVSKLVT	236		
Db	201 CFQTKAATVTKELRESSLNQHACAVMKNFTGTFQAITVTKLSQKTKVNFTEIKLVL	260		
QY	237 DLTKVHTECCGDLLECCADDRADLAKYICENQDSISKLKECEKPLLEKSHCIAEVEND	296		
Db	261 DVAHVHEHCCRGDVLDCLODGEKIMSYICSQDTLSNKITECCCKLTJERGCQIIHAEND	320		
QY	297 EMPADLPSLAADFVESKDVCKNAEAKDVELGMELVEYAPRHPDYSVLLLR;AKTYETT	356		
Db	321 EKPEGLSPNLRPLUGDRDFNQFSGGEKNIFLAGFVHEYSRRHPQALAVSVILRVAKGTQEL	380		
QY	357 LEKCAAAADPHECYAKVDFEFKPLVEPQNLIKONCELFQELGEYAFQNALLVRYTKVP	416		
Db	381 LEKCFOTENPLECDKGEEELQYIQESQALAKRSCLGFKLGEYVLQNAFLVAYTKAP	440		
QY	417 QVSTPTLVEVSRLNGVSKCKKHPKAKRMPCABDYLSVVNLQCVLHEKTPVSDRVTKC	476		
Db	441 QLTSSSELMAITRKMAATAATCCQLSEDKLACGEAGADIIGHLCIRHEMTPTVNPVGQC	500		
QY	477 CTESLVNRRPCFSALEVDETYVPKFNAAETTFHADICTLSEKERQIKKOTALVELVKHK	536		
Db	501 CTSYANRRPCFSSLVVDETYVPPAFSDDDKFIHKDLCCQAGQVALQTMKGEFLINLVKQK	560		
QY	537 PKATKEQLKAVMDFAAFVEKCKKADDKETCFAEEGKGLVAASQAALGL	585		
Db	561 PQITEEQLEAVIADFSGLLEKCCQCGQEQVEYCFAEEGQKLISKTRALGV	609		

Search completed: October 27, 2003, 15:33:07
 Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:29:19 ; Search time 106 Seconds
(without alignments)
1424.158 Million cell updates/sec

Title: US-09-832-929-18

Perfect score: 3103

Sequence: 1 DAHKEVAHFRKDLGEENFK.....TCFAEGKKLVAAQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rhodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2504	80.7	608	5	Q95VB7	Q95vb7 schistosoma
2	2374	76.5	608	11	Q8C7H3	Q8c7h3 mus musculus
3	2336	75.3	576	11	Q8C7C7	Q8c7c7 mus musculus
4	1865.5	60.1	396	4	Q8IUU7	Q8iuk7 homo sapien
5	1295.5	41.7	527	13	Q8JIA9	Q8jia9 sphenodor: p
6	1242	40.0	609	6	Q8MTU5	Q8mju5 canis famiil
7	1242	40.0	626	13	Q8UW05	Q8uw05 ambystoma m
8	1218.5	39.3	610	6	Q8WJ76	Q8wj76 sus scrofa
9	1087	35.0	624	13	Q8UW06	Q8uw06 ambystoma t
10	1082	34.9	605	11	Q8SK65	Q8bk65 mus musculus
11	1076	34.7	605	11	Q8SK65	Q8bk65 mus musculus
12	1045	33.7	400	13	Q8JIA7	Q8jia7 sphenodor: p
13	955	30.8	603	13	Q9YGH6	Q9ygh6 rana shqipe
14	928.5	29.9	614	13	Q91134	Q91134 raja naja (
15	888	28.6	406	13	Q8JIA8	Q8jia8 hoplodactyl
16	739	23.8	205	11	Q8CG74	Q8cg74 mus musculus

17	713	23.0	417	11	Q8R0J9	Q8r0j9 mus musculu
18	376.5	12.1	484	13	Q9W6F5	Q9w6f5 gallus gall
19	373	12.0	476	11	Q9CY31	Q9cy31 mus musculu
20	372	12.0	476	11	Q9IXG1	Q9ixg1 mus musculu
21	365	11.8	551	13	O42279	O42279 petromyzon
22	331	10.7	122	13	Q9WZ8	Q9wz8 latus argen
23	290	9.3	123	13	Q9WZ6	Q9wz6 poephila gu
24	264	8.5	135	11	O63205	O63205 rattus norv
25	188	6.1	1723	2	Q9JMX8	Q9jmx8 helicobacte
26	184	5.9	1819	16	Q9ZLV0	Q9zlv0 helicobacte
27	184	5.9	1927	16	O25262	O25262 helicobacte
28	162.5	5.2	44	6	Q9SMC2	Q9smc2 equus cabal
29	162.5	5.2	680	5	Q9V6S8	Q9v6s8 drosophila
30	161	5.2	1079	3	Q96V11	Q96v11 pneumocysti
31	156	5.0	1026	3	O74669	O74669 pneumocysti
32	154.5	5.0	3843	5	Q9USD0	Q9usdc drosophila
33	153.5	4.9	661	5	O8MS79	O8ms79 drosophila
34	153.5	4.9	3843	5	Q9VU94	Q9v94 drosophila
35	153	4.9	62	6	O8MIL1	O8mll1 sus scrofa
36	150	4.8	1065	3	Q01828	Q01828 pneumocysti
37	149	4.8	40	6	O9TRA5	O9tra5 oryctolagus
38	148	4.8	1028	3	O74668	O74668 pneumocysti
39	144.5	4.7	8749	4	O8NFP1	O8nfp1 homo sapien
40	142.5	4.6	2756	10	Q9LJ60	Q9lj60 arabidopsis
41	141.5	4.6	1560	5	O26644	O26644 strongyloce
42	138.5	4.5	1069	3	Q96V12	Q96v12 pneumocysti
43	137.5	4.4	1348	16	O8YK55	O8yk55 anabaena sp
44	135.5	4.4	2841	5	O8MLU9	O8mlu9 drosophila
45	135.5	4.4	2931	5	Q9W2C6	Q9w2c6 drosophila

ALIGNMENTS

RESULT 1

Q95VB7 ID Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
component in murine infection with Schistosoma mansoni";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF418550; AAL08579.1; -
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EABB28E1C66E54 CRC64;

Query Match 80.7%; Score 2504; DB 5; Length 608;

Best Local Similarity 76.3%; Pred No. 1.7e-188;

Matches 445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;

Qy 1 DAHKEVAHFRKDLGEENFKALVIAFAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAE 60

DB 25 DAHKEVAHFRKDLGEHFKGLVIAFSQLOKCPYEEHVKLVNEVTDFAKTCVADESAE 84

Qy 61 NDKSKLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 120

DB 85 NDKSKLHTLFGDKLCAIPTLRDSYGLADCCAKKEPERNECFLOHKDDHNPFPVRPDA 144


```

QY 121 DVMCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAATECCCAKAAACLLP 180
DB 145 EAMCTSPQENAVFTPMGHYLVHVARRRHPYFYAPPELLYAEKYSIAIMTECCGEACRAACITP 204
QY 181 KLDELDRDEGKASSAKORLKASLOKFEGERAFKAWAVARLSORFPKAEFAEVSCLVDTLTK 240
DB 205 KLDGVTKEKALVSSVRQRMKSSMQKFEGERAFKAWAVARLSQTFPNADFAEITKLATDLTK 264
QY 241 VHTTECCGHDLECCADRADLAKYICENQDSISSKLKECCKEPPLLEKSHGICAEVDEMPA 300
DB 265 KLDALKEKALASSVNORLKCSSLQRFQORAFKAWAVARMSQKFPKADFAEITKLATDLTK 264
QY 241 VHTTECCGHDLECCADRADLAKYICENQDSISSKLKECCKEPPLLEKSHGICAEVDEMPA 300
DB 265 LTBECCHGDLLECCADRAELAKYMCENQATISSKLQCCCKPVLKSHCLSEVENDLPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVLLRLAKYETITLTK 360
DB 325 DLPSLAADFVEDKVCNKYAEAKDVFGLTFLEYARRHPDYSVALLRLAKYETITLTK 384
QY 361 CAADAPHECVAKYFDEPKPLVEBPONLIKONCELFOLGEYFQNALVRYTKKAPQVST 420
DB 385 CAADAPHECVAKYFDEPKPLVEBPONLIKONCELFOLGEYFQNALVRYTKKAPQVST 444
QY 421 PTLVEVSRLNGKVGSKCCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 540
DB 505 LVNRRPCFSALEVDETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 564
QY 541 KEOLKAYMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAAL 583
DB 565 GPQRLTVLGEFTAFDCKCAEDKAEACFSDGPKLVASSQAAL 607

```

```

RESULT 2
Q8C7H3 PRELIMINARY; PRT; 608 AA.
AC Q8C7H3: 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Albumin 1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
DR EMBL; AK050248; BAC34145.1;
SQ SEQUENCE 608 AA; 68722 MW; 292F600BED3A61B4 CRC64;
Query Match 76.5%; Score 2374; DB 11; Length 608;
Best Local Similarity 72.2%; Pred. No. 2.7e-178;
Matches 421; Conservative 80; Mismatches 82; Indels 0; Gaps 0;

```

```

QY 1 DAHSEVARRFKDLEENFKALVLIATAFYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 EAHKSEIARNYDLGEOHFGLVLIATFSYLOKCSYDEHAKLVQEVTDFAKTCVADESAA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 144
QY 121 DVMCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAATECCCAKAAACLLP 180
DB 145 EAMCTSPQENAVFTPMGHYLVHVARRRHPYFYAPPELLYAEKYSIAIMTECCGEACRAACITP 204

```

```

QY 181 KLDELDRDEGKASSAKORLKASLOKFEGERAFKAWAVARLSORFPKAEFAEVSCLVDTLTK 240
DB 205 KLDGVTKEKALVSSVRQRMKSSMQKFEGERAFKAWAVARLSQTFPNADFAEITKLATDLTK 264
QY 241 VHTTECCGHDLECCADRADLAKYICENQDSISSKLKECCKEPPLLEKSHGICAEVDEMPA 300
DB 265 VNECCCHGDLLECCADRAELAKYMCENQATISSKLQCCCKPVLKSHCLSEVENDLPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVLLRLAKYETITLTK 360
DB 325 DLPSLAADFVEDKVCNKYAEAKDVFGLTFLEYARRHPDYSVALLRLAKYETITLTK 384
QY 361 CAADAPHECVAKYFDEPKPLVEBPONLIKONCELFOLGEYFQNALVRYTKKAPQVST 420
DB 385 CAADAPHECVAKYFDEPKPLVEBPONLIKONCELFOLGEYFQNALVRYTKKAPQVST 444
QY 421 PTLVEVSRLNGKVGSKCCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 540
DB 505 LVNRRPCFSALEVDETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 564
QY 541 KEOLKAYMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAAL 583
DB 565 AEQLKAYMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAAL 607

```

```

RESULT 3
Q8C7C7 PRELIMINARY; PRT; 576 AA.
AC Q8C7C7: 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Albumin 1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
DR EMBL; AK050644; BAC34360.1;
FT NON_TER 1
SQ SEQUENCE 576 AA; 65002 MW; F85733BE99AE37F04 CRC64;
Query Match 75.3%; Score 2336; DB 11; Length 576;
Best Local Similarity 72.2%; Pred. No. 2.5e-175;
Matches 415; Conservative 79; Mismatches 81; Indels 0; Gaps 0;

```

```

QY 9 HRFKDLGEENFKALVLIATAFYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 68
DB 1 NRYNDLGEQHFGLVLIATFSYLOKCSYDEHAKLVQEVTDFAKTCVADESAAE 60
QY 69 LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEVDMCTAFH 128
DB 61 LFGDKLCAIPNLRNRYNGELADCCAKQEPERNECFLOHKDDNPRLVRPEVDMCTAFH 120
QY 129 DNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAATECCCAKAAACLLPKLDELRE 188
DB 121 ENPTFMGHYLVHVARRRHPYFYAPPELLYAEKYSIAIMTECCGEACRAACITP 180
QY 189 GKASSAKORLKASLOKFEGERAFKAWAVARLSORFPKAEFAEVSCLVDTLTKVHTCECHG 248
DB 181 ALVSSVQRMKSSMQKFEGERAFKAWAVARLSQTFPNADFAEITKLATDLTKVHTCECHG 240

```



```
QY 249 DLLCADDADLAKYICNODSISSKLEKCEKPLEKSHCIAEVENDEMPADLPSAA 308
DB 241 DLLCADDADLAKYICNODSISSKLEKCEKPLEKSHCIAEVENDEMPADLPSAA 300
QY 309 FVEKQVCKYVABAKDVLGFLMFLYVARRHDPYVSVLLRLAKYETTTLEKCCAAAPHE 368
DB 301 FVEQEVCKYVABAKDVLGFLMFLYVARRHDPYVSVLLRLAKYETTTLEKCCAAAPPA 360
QY 369 CYAKVDFEFKPLVEBPQNLKONCELFQOLGEYFQNALVRYTKKVPQVSTPTLVEVSR 428
DB 361 CYGTVLAEFQPLVEBPQNLKONCELFQOLGEYFQNALVRYTKKVPQVSTPTLVEAAR 420
QY 429 NLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRPCF 488
DB 421 NLGRVGTCKCTLPEDORLPCVEDYLSALNRCVLLHEKTPVSEHVTKCCSSGLVRRPCF 480
QY 489 SALEVDETYVPKEFNAETFTHADICTLSEKERQIKKQTALVELVGHKPKATKEQJXVM 548
DB 481 SALTVDETYVPKEFAETFTFHSOICTLPEKEKQIKKQTAJAEVLVGHKPKATBQJKTVM 540
QY 549 DDAFAVEKCKKADDDKETCFABEKGKLVAAASQAAL 583
DB 541 DDAFAVLOCTCCAAADKOTCFSTEGENLVTRCKDAL 575

RESULT 4
Q81UK7 PRELIMINARY; PRT; 396 AA.
AC Q81UK7
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969; AAH35969.1; -.
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;
```

```
Query Match 60.1%; Score 1865.5; DB 4; Length 396;
Best Local Similarity 63.6%; Pred. No. 1.6e-138;
Matches 372; Conservative 0; Mismatches 0; Indels 213; Gaps 1;

QY 1 DAHKEVAHRRFKDLGEENFKALVLIAFAQYLQOCQPFEDHVKLVNEVTEFAKTCVADESA 60
DB 25 DAHKEVAHRRFKDLGEENFKALVLIAFAQYLQOCQPFEDHVKLVNEVTEFAKTCVADESA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKGEPEKNECFLOHKDNDNPLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKGEPEKNECFLOHKDNDNPLRLVRPEV 144
QY 121 DVMCTAFHNDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAACLLP 180
DB 145 DVMCTAFHNDNEETFLUKYV----- 163
QY 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRFPKAEFVSKLVTDLT 240
DB 164 ----- 163
QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
DB 164 ----- 163
QY 301 DLPSLAADFVSKQVCKYVABAKDVLGFLMFLYVARRHDPYVSVLLRLAKYETTTLEK 360
DB 164 ----- 163
QY 361 DLPSLAADFVSKQVCKYVABAKDVLGFLMFLYVARRHDPYVSVLLRLAKYETTTLEK 360
DB 164 ----- 163
QY 429 NLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRPCF 488
DB 421 NLGRVGTCKCTLPEDORLPCVEDYLSALNRCVLLHEKTPVSEHVTKCCSSGLVRRPCF 480
QY 489 SALEVDETYVPKEFNAETFTHADICTLSEKERQIKKQTALVELVGHKPKATKEQJXVM 548
DB 481 SALTVDETYVPKEFAETFTFHSOICTLPEKEKQIKKQTAJAEVLVGHKPKATBQJKTVM 540
QY 549 DDAFAVEKCKKADDDKETCFABEKGKLVAAASQAAL 583
DB 541 DDAFAVLOCTCCAAADKOTCFSTEGENLVTRCKDAL 575
```

```
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLKONCELFQOLGEYFQNALVRYTKKVPQVST 420
DB 172 CAAADPHECYAKVDFEFKPLVEBPQNLKONCELFQOLGEYFQNALVRYTKKVPQVST 23;
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
DB 232 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 291
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTHADICTLSEKERQIKKQTALVELVGHKPKAT 540
DB 292 LVNRRPCFSALEVDETYVPKEFNAETFTHADICTLSEKERQIKKQTALVELVGHKPKAT 351
QY 541 KEOLKAVMDDDFAAFVEKCKCKADDDKETCFABEKGKLVAAASQAALGL 585
DB 352 KEOLKAVMDDDFAAFVEKCKCKADDDKETCFABEKGKLVAAASQAALGL 396

RESULT 5
Q8JIA9 PRELIMINARY; PRT; 527 AA.
AC Q8JIA9
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Serum albumin (Fragment).
OS Sphenodon punctatus (Hatteria) (Tuatara).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
RP SEQUENCE FROM N.A.
RA Metcalf V.J., Brennan S.O., George P.M., Chambers G.K.;
RT "Partial coding sequence for Sphenodon punctatus 68 kDa albumin.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375971; AAM46104.1; -.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON TER 1
SQ SEQUENCE 527 AA; 59711 MW; C62B799B387F5929 CRC64;
```

```
Query Match 41.7%; Score 1295.5; DB 13; Length 527;
Best Local Similarity 45.9%; Pred. No. 1.4e-93;
Matches 240; Conservative 92; Mismatches 188; Indels 3; Gaps 3;

QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKGEPEKNECFLOHKDNDNPLRLVRPEV 120
DB 5 CLKSLDTIFDEICHEEGEFAAKY-DLAAACCAKAEVERKECLLAHKQATPGFIPAFORPGI 63
QY 121 DVMCTAFHNDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAACLLP 180
DB 64 EVSKCLYQDDRLTLGNYIYEVARRHPYLQVPVFATASLYDEALKTCCTADKATCFHP 123
QY 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRFPKAEFVSKLVTDLT 240
DB 124 RIPPLLEYLKMNGIOENTCGILKKEGERTLAKTLVQMSQKFPKADFATINKLVEDITH 183
QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
DB 184 MHTECCRGTLLECLRDREALTEYTCCHKDAISKLPCTCCEKSVLERGECIVRLNDDKFA 243
QY 301 DLPSLAADFVSKQVCKYVABAKDVLGFLMFLYVARRHDPYVSVLLRLAKYETTTLEK 360
DB 244 DLSERIAEYIEDPHVCDHLAKEQDAFLAKFLYEYSRRHPSELSTQILLGVGKGQVOLLERC 303
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLKONCELFQOLGEYFQNALVRYTKKVPQVST 420
DB 304 CKTDNPPCYGQAEADLKKHIAQFQELVQONCDLYNTLGGYLFHNNALLIRYTKRMPQLTS 363
```

```

QY 421 PT-VEVSRNLGKYGSKCCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTESLVNRRPCFSALEVD 480
DB 364 EELFYTR-ITKAASRCEVSUVKLPCTEGYVDFVLGQCCQHRSNNVNVQCCSN 422
QY 481 LVNRRPCFSALEVDYVVPKFNATFTTHADICTLSEKERQIKKOTALVELVGHKPKAT 540
DB 423 YALASLCITSLGSGDEKVPTEFSADLFTFHEDLCHAAQDKLOERKQOMIVNLVGHKPNIT 482
QY 541 KEOLKAYMDDFAFVEKCCKADDKETCFABEGKLVAAASQAL 583
DB 483 KEQLQTVFGFTTKWTEKCKCAEDHEACFGBEGPKLVAESQAL 525

RESULT 6
Q8MJU5 PRELIMINARY; PRT; 609 AA.
AC Q8MJU5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Alpha-fetoprotein.
GN AFP.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Furuichi M., Neo S., Hisasue M., Tsuchiya R., Watanabe M.,
RA Hashizaki K., Hisanatsu S., Yamada T.;
RT "Canine alpha-fetoprotein cDNA.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089789; BAC07513.1; -.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; P202486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 609 AA; 68782 MW; BE4B9250C5AF2AF0 CRC64;

```

```

Query Match 40.0%; Score 1242; DB 6; Length 609;
Best Local Similarity 40.5%; Pred. No. 2.7e-89;
Matches 231; Conservative 113; Mismatches 225; Indels 2; Gaps 2;

QY 16 EENFKALVLAFAQYLOQCPEFDHVKLVNVEFTFAKTCVADSAENKDSLHFLFGDKLC 75
DB 40 EENLVDLATIFFAQFVQEAQYKEYSKVKDILTVEIKSTGSEQGGCLENLQAPAELEIC 99
QY 76 TVATLRETYGEMADCCAKQEPERNECELOHKD-DNPNLPRLVPEVDVDMCTAFHDNEETP 134
DB 100 HEKISEIKYG-LADCCSQREERHNCFLAHKKAAPSPFPQVAEVTCKAYEENRDMF 158
QY 135 LKRYLYEIARRHPFYAPPELLFFAKRYKAAFTCCQAADKAAAGLLPKLDELDEGRKASSA 194
DB 159 MNRVYIEIARRHPFLYAPTILSLAHHYKGIPLCCRAENAVECFQTKSLTKELRESSJ 218
QY 195 KQRLKASLQKGBERAFKANAVARLSORFPKAPFAEVSKVLVTLTKVHTCECHGLLECA 254
DB 219 LNQHICAMRNFGPRFTFRATITVTKLSQKSKANFTETQKLVLDVAIHBECCRGNVLECL 278
QY 255 DDRADLAKYICENODSISSKLKECCPELLEKSHCIAEVDENKPADLPSLAADRVESKD 314
DB 279 QDGEKIMSYICQODILUSSKIADCCKLPILLEGQCIHAENDKPGELSPNLRFLERD 338
QY 315 VKKNYAEAKVFLGMFLYIARRHPDYSVVLLRLAKTYETTLLEKCAAADPHECVAKVF 374
DB 339 FNQFSREKDLFMAFTYIYSRRHTKLVAPVWLVAKYGELEKESQSENPLECODKGE 398
QY 375 DEFKPLVEEPONLTKONCELFPEQLGEYFQNALLVRYTKVPQVSTFTLVEVSRNLGKYG 434
DB 399 BELEKYIQESQALAKRSGLFQKLGYYLQNAFLVAVTKKAPQLTPPELMVAFTRKNATAA 458

```

```

QY 435 SKCKKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTESLVNRRPCFSALEVD 494
DB 459 ATCCQLSEDLQACGEGADLLIIGQICIRHEETPINPGVGQCCSSSYANRRPCFSSLVVD 518
QY 495 ETVVPKEFNAETFTTHADICTLSEKERQIKKOTALVELVGHKPKATKEQLKAYMDDFAAF 554
DB 519 ETVIPSPFSADKRFIFHKDLCOAGVALQTMKQFLNLVKQKQPIITEEQLEAVIADFSGL 578
QY 555 VEKCKKADDKETCFABEGKLVAAASQALGL 585
DB 579 LEKCCQCEQAEACFELEGPKLISKTRAAALGV 609

RESULT 7
Q8UW05 PRELIMINARY; PRT; 626 AA.
AC Q8UW05
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Serum albumin precursor.
GN AIB.
OS Ambystoma maculatum (spotted salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=43114;
RN [1]
RP SEQUENCE FROM N.A.
RA Haverfield E.V., Uzzell T., Spoisky C.M., Bazarstesen B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RT Ambystoma texanum.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217183; AAL56646.1; -.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 626 SERUM ALBUMIN.
SQ SEQUENCE 626 AA; 70677 MW; 9D66F57F174AC23F CRC64;

Query Match 40.0%; Score 1242; DB 13; Length 626;
Best Local Similarity 40.1%; Pred. No. 2.8e-89;
Matches 237; Conservative 110; Mismatches 238; Indels 6; Gaps 3;

QY 1 DAHKSEVAHRFKDL-----GEENFKALVLAFAQYLOQCPEFDHVKLVNVEFTFAKTCVAD 56
DB 28 EGHVDNPPHLIGDLTPMIGVDNSKGLVLAASQMLPLCPVEEHLQORVEDVMQIADLCAG 67
QY 57 ESAENCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECELOHKDDN-PNLPRJ 115
DB 86 ARHANCASPMWIIIDELCKKPPNAEKYPHQECCKEDPERHKCFVEHQNAHEELTKY 147
QY 116 VRPEVDVMCTAFHDNEETFLKKYLYEIARRHPFYAPPELLFFAKRYKAAFTCCQAADKA 175
DB 148 VRPAPEQICKDHAENRGPLLARYIFMLAIGHPHMYIPAILGFAQRFDGIVSHCKDVETA 207
QY 176 A-CLLPKLDELDEGRKASSAKQRLKASLQKGBERAFKANAVARLSORFPKAPFAEVSKL 234
DB 208 GQCENDKMPHKEQVEYVCAQKNGCYILQDFKERALTATYKAVQASOKFPPLASFENVQI 267
QY 235 VTDLTKVHTCECHGLLECCADDDRADLAKYICENODSISSKLKECCPELLEKSHCIAEVE 294
DB 268 VPDTVHUHQTCGGDMACMLERKWTAKICEKKDELATHLKECCDKPLERSACILRP 327
QY 295 NDEMPADLPSLAADRVESKDVCKNVAEKDVFLGMFLYIARRHPDYSVVLLRLAKTYE 354
DB 328 NDQKPADLSPKVPHYIDDPEVCKLYTEGGDTFNGRFLYECARRHQDYSPEMLLRMSSGYE 387

```

```

Qy 355 TTLEKCCAAADPHCEYAKVDEFEKPEVEEPONLIKQNCCELFEQLGEYKFKQALLVRYTKK 414
Db 388 EFLKRCCEAAGHNECCATESLKKEIESSVTLTKNGCAJDKLKSYLQKLLIFKVIAR 447
Qy 415 VPQVSTPTLVEVSRLNUGKVGSKCKCHPEAKMPCAEADYLSVVJNQLCVLHKEIFVSRRVT 474
Db 448 MPALSEQSLRITKMTTIGEKCKCHRPEDQOMTCSEGGGIVFGICMKQKTTTPNEKVA 507
Qy 475 KCCTESLVRNRPCCPSALEVEDETVVPKFNENAEFTTFHADICTLSEKERQIKKQALVELVK 534
Db 508 QCCSHLSLSQPCPSALEVEDETVVPKFNENAEFTTFHADICTLSEKERQIKKQALVELVK 567
Qy 535 HKPKATKEQLKAVMDDEFAAFVEKCKKADKCTCFABEGKKLVAASQAALGL 585
Db 568 QYPHMTDEQLKTCVNVFVPMVDQCCNHNNECFALGAKLIDACKALJAV 618

RESULT 8
Q8MJ76 PRELIMINARY; PRT; 610 AA.
AC Q8MJ76;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Alpha-fetoprotein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim J.G., Noneman D., Vallet J.L., Christenson R.K.;
RT "Mapping of the porcine alpha-fetoprotein (AFP) gene to SSC8.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF517770; AAM66710.1; -.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 610 AA; 68624 MW; C985BEAD44963D5E CRC64;

Query Match 39.3%; Score 1218.5; DB 6; Length 610;
Best Local Similarity 39.7%; Pred. No. 1.9e-87;
Matches 227; Conservative 117; Mismatches 225; Indels 3; Gaps 3;

Qy 16 BENFKALVLIAPAOYLQOCPEFDHVKLVNEVTEFAKTCVADESANCDKSLHTLFGDKLC 75
Db 40 EMNLVDLATITFAQFQVQATYKEVNMQKDVLTVEKSTGSEQAPAGCLENQVSVFLEIC 99
Qy 76 TVATLRETYGEKADCCAKQEPERNECFLOHKDDNP-NLPRLVRPEVDVWCTAFHNDNETF 134
Db 100 HEEEPKPYG-LSHCCSQSGEERHNCFLARKKAAPAS:PPFQVPPVTSKAYEENREL 158
Qy 135 LKKYLYETARHPYFAPPELLFFAKRYKAAFTCCQADKAAACLPKLDELDECKASSA 194
Db 159 MTRYIYETARRHPFLYATITLSLAAYQDKITPPCCKAENAVECFQTKAASITKELRESSL 218
Qy 195 KQRLKASLQKFEGERAFKAWAVARLSQRFPKAEFAEVSKLVTDJTKVHTCHGDLLECA 254
Db 219 LNHQWCTVNRQFGATFRATITVKLSQFPKANFTIQKLVLDVAHIHEECRCGNVLECL 278
Qy 255 DRRADLAKYICENQDSISGKLKECKECP-LLEKSHCIAEVENDEMPADLPSLAADFVESK 313
Db 279 QDAERWVSVCQDQLSSKIAECCKLP-TTLELQGCIIHAENDDKPEGLSPNLNRFLEGR 338
Qy 314 DVCKNYAEAKDVFLGMFLYEVARRHPDVSULLRLAKTYETTTLEKCCAAADPHCEYAKV 373
Db 339 DFNQLSSREKLSMARFTYEYSRRHPKLAVPVILRVAKGYQELLEKCSQSENPCECDKG 398
Qy 374 PDEFKPLVEEPONLIKQNCCELFEQLGEYKFKQALLVRYTKKVPQYSTPTLVEVSRLNUGK 433

```

```

Db 399 EEELEKVIQESQALAKRSQGLFQKLGCEYVQNAFLVAYTKKAPQLTPPELMALTRQWATT 458
Qy 434 GSKCCXHPHAKMPCAEADYLSVVJNQLCVLHKEFTPVSDRVTKCCTESLVRNRPCCPSALEV 493
Db 459 GAACCHLSDEQLACGEAADLIIGQLCIRHBEEMPINPGVGGCCTSSYANRRPCCSSVLV 518
Qy 494 DETYVVKFNENAEFTTFHADICTLSEKERQIKKQALVELVKHKKPKATKEQLKAVMDDEFAA 553
Db 519 DETYVVPFPPSDDKFIHKLQCCQAGVALQTMKQOFLINLVKQKQITQTEQLEAVIADFSG 578
Qy 554 FVEKCKKADKCTCFABEGKKLVAASQAALGL 585
Db 579 LLEKCCQGEQEVCFABEGPALISKTRASLGV 610

RESULT 9
Q8UW06 PRELIMINARY; PRT; 624 AA.
AC Q8UW06;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ambystoma texanum (Smallmouth salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8304;
RN [1]
RP SEQUENCE FROM N.A.
RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazariseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
R: Ambystoma texanum.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217182; AAL56645.1; -.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 1.
KW Signal.
FT SIGNAL. 1 24 POTENTIAL.
FT CHAIN 25 624 SERUM ALBUMIN.
SQ SEQUENCE 624 AA; 70321 MW; DE08533BF4953BF7 CRC64;

Query Match 35.0%; Score 1087; DB 13; Length 624;
Best Local Similarity 37.7%; Pred. No. 4.3e-77;
Matches 214; Conservative 102; Mismatches 240; Indels 12; Gaps 5;

Qy 14 LGSENFKALVLIAPAOYLQOCPEFDHVKLVNEVTEFAKTCVADESANCDKSLHTLFGDK 73
Db 45 IGVEHAKALAMALFQSMLSKCPHHEQVQVRVNVMDIADLCRSGAKGDCGKSVMTILNE 104
Qy 74 LCTVATLRETYGEKADCCAKQEPERNECFLOHKDDNP-NLPRLVRPEVDVWCTAFHNDNEE 132
Db 105 ICKTPENPEKYPFHGECCKEDPERHAKCFIEHSTQPKERTKTEYVKPSPEQICKHAEHRD 164
Qy 133 TFLKKYLYETARHPYFAPPELLFFAKRYKAAFTCCQADKAAACLPKLDELDECKASSA 191
Db 165 EFLGHYIHKVASSHTTMYPPAILSLFTLHPDGIHSHCKDEATVQCCLSEKMPAHKEVEH 224
Qy 192 SSAQRKLKASLQKFEGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTQVHTCHGDLLE 251
Db 225 VCAVQKHNCVILQNFENRALRASKAAHACSKFPFASFPNVQRLTDGIVLHLHQTCGGDMM 284
Qy 252 ECADRRADLAKYICENQDSISGKLKECKECP-LLEKSHCIAEVENDEMPADLPSLAADFVE 311
Db 285 ACMAERMKLTQTCEK-----KKCEKPVLERSECVRLPNDEKPADLSPEVRYFD 336
Qy 312 SKDCKNYAEAKDVFLGMFLYEVARRHPDVSULLRLAKTYETTTLEKCCAAADPHCEYAKV 371

```

```

Db      337 DPEVCKRFXEGDAPMGRFLCDYAK:HPHSRAFLNLRASG:EKAYKTCAGAEANECEIA 396
QY      372 KVDFEFKPLVESQNLKONCELFEGQGEYKFNALVRYTKKVQVSTPTTVVSVRNLG 431
Db      397 KBETLRIHEIASKTKLKTCGALEKLGPHYHFQINIMVRYTGILPRSSDAFLYITKTL 456
QY      432 KVGSKCKKHPKARMPCAEDYLSVLNQLCVLHETPV-SDRVTKCCTESLVNRRPCFSA 490
Db      457 NCQKCKKLPEDQMPCEBGLGNVFAIC-QNKTPFENEKLAHCKKSLSFTTTPCFAA 515
QY      491 LEYDETVPKEFNAETFTFHADICTLSEKEROIKKQTAIVEVPHKPKATKEQLKAVMD 550
Db      516 LTVDETYVPAPVTAESFNENDEFTSEADLOAKQOTFLMLHVRTHPKITQSOVXTISEK 575
QY      551 FAAPVEKCKCKADDKETCFABEGKKLVAA 578
Db      576 FLAMGCGCKCKADQORNECFATEGAKLVEA 603

```

RESULT 10

```

Q8BK65 PRELIMINARY; PRT; 605 AA.
ID Q8BK65
AC Q8BK65;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Alpha-fetoprotein precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; AK076053; BAC36150.1; -.
SQ SEQUENCE 605 AA; 67322 MW; 048B74A8B01EA4B CRC64;

```

```

Query Match      34.9%; Score 1082; DB 11; Length 605;
Best Local Similarity 35.7%; Pred. No. 1e-76;
Matches 204; Conservative 120; Mismatches 241; Indels 6; Gaps 3;

```

```

QY      16 EENFKALVLAFAYLQOCCPFEDHVKLVNEVTEFAKTCVADESAENCDSLHTLFQDKLC 75
Db      40 EKNVLSIATITFTQVPPEATEEE----VNKMTSDVLAAMKKNKSGDGLCSQSVFLDEIC 95
QY      76 TVATLRETYGEMADCCAKQBPBERNECFLOHKDNP-NLPLRVPEVDVMTAFHDNEETF 134
Db      96 HETELSNKYG-LSGCCSQSGVERHQCLARKKKTAPASVPPFPPEPAESKAHEENRAVF 154
QY      135 LKXYLYIARHPHYFVYAPPELLFFAKRYKAAFTCCQAAKAAACLLPKLDELDEGKASSA 194
Db      96 HETELSNKYG-LSGCCSQSGVERHQCLARKKKTAPASVPPFPPEPAESKAHEENRAVF 154
QY      135 LKXYLYIARHPHYFVYAPPELLFFAKRYKAAFTCCQAAKAAACLLPKLDELDEGKASSA 194
Db      155 MNRFIYVSRNPFMYAPAILSLAAQYDKVLVLAACKADNKEECFQTKRASIAREGSM 214
QY      195 KQRLKASLOKFGRAFAPKAWARLSQRPFKPAFAEVSKLVTLTKVHTCCGGDLLECA 254
Db      155 MNRFIYVSRNPFMYAPAILSLAAQYDKVLVLAACKADNKEECFQTKRASIAREGSM 214
QY      195 KQRLKASLOKFGRAFAPKAWARLSQRPFKPAFAEVSKLVTLTKVHTCCGGDLLECA 254
Db      215 LNEHVCVIRKFGSRNIQATTIKLSQKLTAEANFTIQKLALDVLIHBECCGNSLECL 274
QY      255 DDRADLAKYICENODSISSKLKECCPEKPLEKSHCIAEVENDEMPADLPSLAADFVESKD 314
Db      215 LNEHVCVIRKFGSRNIQATTIKLSQKLTAEANFTIQKLALDVLIHBECCGNSLECL 274
QY      255 DDRADLAKYICENODSISSKLKECCPEKPLEKSHCIAEVENDEMPADLPSLAADFVESKD 314
Db      275 ODGEKWTYICSQNLISSKIAIECCKLPYQLGFCIIHAENGVKPGLSINPSQFLGDRN 334
QY      315 VCKNYAEAKDVLGMFLYEARHPDYVSVLLLRKATYETTITLLEKCCAAADPHECVAKVF 374
Db      275 ODGEKWTYICSQNLISSKIAIECCKLPYQLGFCIIHAENGVKPGLSINPSQFLGDRN 334
QY      315 VCKNYAEAKDVLGMFLYEARHPDYVSVLLLRKATYETTITLLEKCCAAADPHECVAKVF 374
Db      335 FAQFSSEKIMFMASFHEYSRTHPNLPVSVILRIAKTYOEIILEKCSQSNLPGCCQDNLE 394
QY      375 DEFKPLVEEONLIKONCELFEGQGEYKFNALVRYTKKVQVSTPTTVVSVRNLGKVG 434
Db      335 FAQFSSEKIMFMASFHEYSRTHPNLPVSVILRIAKTYOEIILEKCSQSNLPGCCQDNLE 394
QY      375 DEFKPLVEEONLIKONCELFEGQGEYKFNALVRYTKKVQVSTPTTVVSVRNLGKVG 434

```

```

Db      395 BELQKHIEESQALSQSCALYQTLGDYKLNFLICYTRKAPOLTSABLIDLTGKMWVSA 454
QY      435 SKCCXGHPKARMPCAEDYLSVLNQLCVLHETPVSDRVTKCCTESLVNRRPCFSAJEVD 494
Db      455 STCCQLSEBKWSGCGENADIFTHLCIRNEASPVNSG:SHCCNSSYSNRRLCITTSFLRD 514
QY      495 ETVVPKEFNAETFTFHADICTLSEKEROIKKQTAIVEVPHKPKATKEQLKAVMD3FAAF 554
Db      515 ETVAPPFSSEKXFIHKLDCQAHGKALQTMKQELLNLVKQPELVEEQLAAVTAFCGL 574
QY      555 VEKCKKADDKETCFABEGKKLVAA5QAALGL 585
Db      575 LEXCKKAQDOQVCFTEEGPKLJSKTRDALGV 605

```

RESULT 11

```

Q8BK56 PRELIMINARY; PRT; 605 AA.
ID Q8BK56
AC Q8BK56;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Alpha-fetoprotein precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; AK076197; BAC36249.1; -.
SQ SEQUENCE 605 AA; 67409 MW; CF3509A9EC146110 CRC64;

```

```

Query Match      34.7%; Score 1076; DB 11; Length 605;
Best Local Similarity 35.6%; Pred. No. 3.1e-76;
Matches 203; Conservative 120; Mismatches 242; Indels 6; Gaps 3;

```

```

QY      16 EENFKALVLAFAYLQOCCPFEDHVKLVNEVTEFAKTCVADESAENCDSLHTLFQDKLC 75
Db      40 EKNVLSIATITFTQVPPEATEEE----VNKMTSDVLAAMKKNKSGDGLCSQSVFLDEIC 95
QY      76 TVATLRETYGEMADCCAKQBPBERNECFLOHKDNP-NLPLRVPEVDVMTAFHDNEETF 134
Db      96 HETELSNKYG-LSGCCSQSGVERHQCLARKKKTAPASVPPFPPEPAESKAHEENRAVF 154
QY      135 LKXYLYIARHPHYFVYAPPELLFFAKRYKAAFTCCQAAKAAACLLPKLDELDEGKASSA 194
Db      155 MNRFIYVSRNPFMYAPAILSLAAQYDKVLVLAACKADNKEECFQTKRASIAREGSM 214
QY      195 KQRLKASLOKFGRAFAPKAWARLSQRPFKPAFAEVSKLVTLTKVHTCCGGDLLECA 254
Db      215 LNEHVCVIRKFGSRNIQATTIKLSQKLTAEANFTIQKLALDVLIHBECCGNSLECL 274
QY      255 DDRADLAKYICENODSISSKLKECCPEKPLEKSHCIAEVENDEMPADLPSLAADFVESKD 314
Db      275 ODGEKWTYICSQNLISSKIAIECCKLPYQLGFCIIHAENGVKPGLSINPSQFLGDRN 334
QY      315 VCKNYAEAKDVLGMFLYEARHPDYVSVLLLRKATYETTITLLEKCCAAADPHECVAKVF 374
Db      335 FAQFSSEKIMFMASFHEYSRTHPNLPVSVILRIAKTYOEIILEKCSQSNLPGCCQDNLE 394
QY      375 DEFKPLVEEONLIKONCELFEGQGEYKFNALVRYTKKVQVSTPTTVVSVRNLGKVG 434
Db      395 BELQKHIEESQALSQSCALYQTLGDYKLNFLICYTRKAPOLTSABLIDLTGKMWVSA 454
QY      435 SKCCXGHPKARMPCAEDYLSVLNQLCVLHETPVSDRVTKCCTESLVNRRPCFSAJEVD 494

```

```
Db 455 STCCQLSEKSGCGEGMADIFIGHLCIRNEASPVNSGISHCNCSYSSNRRLCITSFLRD 514
Qy 495 ETVYVPEKFNAEFTTHADICTLSEKQKQKTA-VELVGHKPKATKEOLKAVNMDDFAAF 554
Db 515 ETYAPPPSEDKFIPHKOLCOAQKALOTMKOEJLNLVKOKPELTBEQLAAVTADEFSG 574
Qy 555 VEKCKKADDKETCFABEGKKLVAASQAALGL 585
Db 575 LEKCKKAQDOEVCFTTEGPKLISKTRDALGV 605

RESULT 12
Q8JIA7 PRELIMINARY; PRT; 400 AA.
AC Q8JIA7;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE A/B over-sized serum albumin (Fragment).
OS Sphenodon punctatus (Hatteria) (Tuatara).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
RP SEQUENCE FROM N.A.
RA Metcalf V.J., Brennan S.O., George P.M., Chambers G.K.;
RA "Partial mRNA sequence for tuatara A/B serum albumin.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF375973; AAM46106.1; -.
DR Pfam; PF00264; Serum_albumin.
DR InterPro; IPR000264; Serum_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
FT NON_TER 1
FT CHAIN 1
SQ SEQUENCE 400 AA; 45715 MW; 8DE20609657CF753 CRC64;

Query Match 33.7%; Score 1045; DB 13; Length 400;
Best Local Similarity 47.7%; Pred. No. 5.1e-74;
Matches 186; Conservative 80; Mismatches 124; Indels 0; Gaps 0;

Qy 196 ORLKCAQLQKGFGERAFKAWAVARLSQRFPKAEVSKLVTLDTLVKHTCCGDLLECAD 255
Db 3 EKHSCGFLKSGERAFQANKVLQSQKPKAPFEEIHKVLTATKLQKECCGDMIECLD 62
Qy 256 BRADLAKIYICENQDSISSKKECCCKPLLEKSHCIAEVENDEMPADLPSLAADFVSKV 325
Db 63 DRVEVMAYICSAQVAFSSKIKDCCKEPIVDREVCIQADLDKPADLPSTAGQV-ESTEV 122
Qy 316 CKNYAEAKDVFLGMPF-YEVARRHPDYSVVLLRLAKTYETTLKCCAAADPECVKVD 375
Db 123 KHYEGEKDVFLAHFYEYRRHPHFSSQMLURTKGQYQDILKCKTENPECVGKAGE 182
Qy 376 EFKPLVEEPQNLIKONCELFQGLGYKFNALLVRYTKKVPQVSTFTLVEVSRNLGKYS 435
Db 183 ELARHIQESQELKTHCSFYTSQGPQKQVYVYTKMPQLPAEBLIEISKLTGTVGV 242
Qy 436 KCKHPEAKRMPCAEEDYSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRFCPSALEYDE 495
Db 243 KCCPLSEDKRLSCSEKLSMVLFEICRQHEASPVNNHVTHCCTDSYSEMRPCTTKLGVD 302
Qy 496 TVVPKEFNAEFTTHADICTLSEKQIKQKOTALVELVGHKPKATKEOLKAVMDDFAAFV 555
Db 303 SYVPPFCFSTFLFDEQLCTAPEEARLKKQLTFLVKLTQLKFCIDEQELKKLVTDYHAME 362
Qy 556 EXCKKADDKETCFABEGKKLVAASQAALGL 585
Db 363 EKCCQAKNQECCFSTEGEKLTOEGKALLGV 392
```

RESULT 13

```
Q9YGH6 PRELIMINARY; PRT; 603 AA.
AC Q9YGH6;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Rana shqiperica.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBI_TaxID=44326;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Bushat; TISSUE=Liver;
RA Uzzell T., Hotz H.;
RA "Albumin cDNA sequence of Rana shqiperica: evolutionary changes in
RA frog albumins.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RL EMBL; U0452; AAD09358.1; -.
DR HSP; P02768; IE7B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Signal.
FT NON_TER 1
FT SIGNAL <1
FT CHAIN 24 603
FT CHAIN 24 603
SQ SEQUENCE 603 AA; 69293 MW; 340D3723FA010C99 CRC64;

Query Match 30.8%; Score 955; DB 13; Length 603;
Best Local Similarity 33.8%; Pred. No. 9.9e-67;
Matches 190; Conservative 119; Mismatches 241; Indels 16; Gaps 3;

Qy 12 KDLGEENFKALVLIATAFYLOQCFFPDHVKLVNVEVTEFAKTCVADESAENCCKSJHTLFG 71
Db 37 KAVGKPAVEKLVLMVVAODPEKCSLDEHLKVQAKIIEAVNCCEKHPEAECKKFAELIYH 96
Qy 72 DKLTCTVATLETYGENADCCAKQEPNERNECFLOKODNPNLRLVR-----PEVDVMT 125
Db 97 DIVCKEEDIQLPWTTECCGKAEARTKCFEHR-----VRVEYKINIEBSCK 148
Qy 126 AFHDNEETFLKKYLYEYARRHPFYAPPELLFFAKRYKAAFTCCQADKAAKLLPKLDEL 185
Db 149 EHKEHPORAFSYLSNTAKRHSKLYPPAVLGFAIOYNEITTECCAAEDKACFGERNPQV 208
Qy 186 RDEGKASSAKQRLKCAQLQKGFGERAFKAWAVARLSQRFPKAEVSKLVTLDTLVKHTEC 245
Db 209 KULTNYLEDKHKQKCRVLEKFPERSQALTLVQVSQRFQGNKYDDVEKVETIEIAHLNEDC 268
Qy 246 CHGDLLECADRADLAKYICENQDSISSKKECCCKPLLEKSHCIAEVENDEMPADLP 305
Db 269 CKGDAVECTMERMEATEHICLAKELSSKLSLSDCCAKGVLERTPCILALPNEE--POLPIE 326
Qy 306 AADFEVSKQVKYAEAKDVFLGMPF-YEVARRHPDYSVVLLRLAKTYETTLKCCAAAD 365
Db 327 LKEYEYDEHVCEYQKQKRYLAHFTHDYSRSHQESSQPSCLRVSRGFEMLEKCCASAN 386
Qy 366 PHECYAKVDFEFPKPLVEEPQNLIKONCELFQGLGYKFNALLVRYTKKVPQVSTFTLVE 425
Db 387 SAELCKDAPKLLAALKENEIEISKQNGALEKGFNDFYIQLLVRYFGKMPQVTAOTLVE 446
Qy 426 VSRNLGKVGSKCKHPEAKRMPCAEEDYSVVLNQLCVLHEKTPVSDRVTKCTESLVNRR 485
Db 447 LTGRMAKIGYVCCGLPDNKKQPCAEKLDILLGEMCEREKKTFTINDNVHCCVDSVANRR 506
Qy 486 PCPSALEVDEVTYVPKEFNAEFTTHADICTLSEKQIKQKOTALVELVGHKPKATKEOLK 545
Db 507 PCFTKLGYPANYEAPVMDSESKLHFTADMCKSGADDQLKTKLVLLVFLKMKPTCGKEKLT 566
```


QY 532 LVKHKPRATKEQLKAVNDFFAFAFVEKCKKADDKETCFABEGKKLVAASQAAL 583
Db 354 VWXCKPAITHEQLKAVITDFYGVVEKCKCHGENHEACFLAEGPOLVORTQAAL 405

Search completed: October 27, 2003, 15:35:01
Job time : 108 secs